ISSN 2093-0542

KOREAN SOCIETY FOR INTERNET INFORMATION

The 12th International Conference on Internet

(ICONI 2020)

Dec. 13-16, 2020, Landing Convention Center,

Jeju Shinhwa World, Korea

http://www.iconi.org

Proceedings of ICONI 2020

| Organized by |

Korean Society for Internet Information (KSII)

| Sponsored by |

Message from Honorary Chair

We express our warm welcome to all of attendants of the 12th International Conference on Internet (ICONI) that will be held from December 13~16, 2020, in the Jeju Shinhwa World, Jeju Island, Korea.

Over the past 12 years, ICONI has increased its academic impact and influence in the area of Information Science, Computer Engineering and Internet Technology by bringing together international researchers from academia, industry, and government to exchange novel ideas, explore enabling technologies, discuss innovative designs, and share field trial experiences and lessons learned.

This year, ICONI 2020 again covers broad range of topics which are related with Internet technologies and applications such as Wireless and Sensor Network, Security & Privacy and Multimedia, Image Processing, Intelligent Systems and mobile computing, Internet of Things, wireless communications.
It is firmly believed that each paper to be presented will be a basis for further constructive discussion.

We would like to express our heartfelt gratitude to every one who makes this conference successful. First of all, we are so grateful to all contributors who submit their valuable papers, review the papers, and chair the sessions. We also thank many sponsors for their enormous support. Our special thanks must go to the Keynote Speaker Dr. Dayang N. A. Jawawi from Universiti Teknologi Malaysia, Malaysia.

Last, but not least, we deeply appreciate limitless devotion of Organizing Committee members. Without their contribution, this conference would not be realized.

We truly welcome you all to this beautiful Island and wish you pleasant and joyous stay in Jeju Island!

Dr. Jaehyoun Kim

Honorary Chair of ICONI 2020

Dr. Jaehyoun Kim

Sungkyunkwan University, Rep. of Korea

Message from Conference Chair

On behalf of the organizing committee of ICONI 2020 as well as on my own behalf, I welcome you to the 12th International Conference on Internet (ICONI). In particular, I sincerely thank you for participating in this international conference to present and discuss valuable research results even in the midst of the pandemic.

Dr. Seung Ryul Jeong
Kookmin Univ.,
Rep. of Korea

As you all know, COVID-19 has brought about many changes and challenges. Academic conferences seem to be no exception, as we are forced to hold conferences in a contactless environment. Despite these difficulties, we have been able to successfully plan for ICONI 2020, prioritizing safety measures. More than 203 papers have been submitted in 21 Internet-related fields, including mobile Internet computing, Internet security, Internet of Things (IoT), wireless and sensor networks and multimedia, image processing and Internet application management. Of these, more than 150 papers will be presented both online and offline. Moving forward, ICONI 2020 will be a leading example for non-contact conferences.

I owe a great amount of gratitude to the organizing committee members and all contributors for the extraordinary work they have done in organizing this conference. Without their dedicated support, this event would never have been possible

I wish you all a very pleasant time here in Jeju Island and a productive and successful conference.

Dr. Seung Ryul Jeong

Conference Chair of ICONI 2020

Message from Program Chairs

It is a pleasure to welcome you to the 12th International Conference on Internet, 2020 (ICONI 2020) to be held in the Landing Convention Center, Jeju Shinhwa World, Korea

Over the past 12 years, ICONI has grown to be the major international conference in Internet Technology area. The conference is organized as a set of 21 tracks in the Internet Technology area. ICONI continues the endeavor of high-quality, broad international participation in all areas of Internet Technology.

The successful organization of ICONI has required the talents, dedication and time of many volunteers from Malaysia, USA, China, Saudi Arabia, Pakistan, Nepal, Nigeria, Turkey, Morocco, and Rep. of Korea. Special gratitude and appreciation go to all of organizing committee members as they are primarily responsible for the conference.

We hope that you will find the conference both enjoyable and valuable, and also enjoy the architectural, cultural and natural beauty of Jeju island.

Jeju Island is one of the most popular tourist destinations in Korea and well situated for welcoming the ICONI 2020 participants from around the world.

Dr. Soo-Kyun Kim
Jeju National Univ., Rep. of Korea

Dr. Imran Ghani
Indiana Univ. of Pennsylvania, USA

Dr. Soo-Kyun Kim Dr. Imran Ghani
Program Chairs of ICONI 2020

Keynote Speaker

Keynote Speaker Prof. Dr. Dayang N. A. Jawawi

School of Computing, Faculty of Engineering,
Universiti Teknologi Malaysia, Malaysia

Title: Computational Thinking through Educational Robotics

Abstract of the talk

Robots are becoming a popular educational tool in areas of science and technology for primary and secondary school and in several areas of engineering and technology in universities for teaching several subjects, such as math, computer science, mechanics, technology, electronics, programming, artificial intelligence, and computer vision. Thus, the popularity of Educational robotics (ER) has resulted to the proliferations of similar ER pedagogical tool being developed with slight differences based on its educational context level.

In the recent digital revolution, digital literacy is becoming a key factor in education today. Computational Thinking (CT) is a new digital literacy which is gradually being introduced in school education due to its applicability in daily problem-solving process. ER has been increasingly used as a pedagogical tool to attract students in learning computer programming, and when integrated with CT, they can be used to develop high-order thinking skills. In this talk I will present the challenges in intertwining between CT and ER in teaching computing and solving problem skills.

Realizing the benefits of robots in teaching, in 2008 Universiti Teknologi Malaysia (UTM) has introduced ER in class formal education for computing students and in 2012 ER for Co-Curricular Service Learning (CCSL) program was initiated for the purpose of teaching the problem solving and computer programming for school students. An action research conducted in an outreach program aimed at nurturing CT to school student; and promoting generic skills to university student. Finally, in my talk I will discuss the relevance, impact, and prospects of ER in CT and computing education.

Organizing Committee

| Honorary Chair |

Jaehyoun Kim, Sungkyunkwan University, Rep. of Korea

| Conference Chair |

Seung Ryul Jeong, Kookmin University, Rep. of Korea

| Program Chairs |

Imran Ghani, Indiana University of Pennsylvania, USA

Soo-Kyun Kim, Jeju National University, Rep. of Korea

| Steering Chairs |

Junchul Chun, Kyonggi University, Rep. of Korea

Shafie Abd. Latiff, Universiti Teknologi Malaysia

Mingoo Kang, Hanshin University, Rep. of Korea

| Publication Chairs |

Mohd Adham Bin Isa, Universiti Teknologi Malaysia

Sangmin Lee, CHA University, Rep. of Korea

Ziyang Liu, Kyonggi University, Rep. of Korea

In Seop Na Chosun University, Rep. of Korea

| Workshop Chairs |

Radziah Mohamad, Universiti Teknologi Malaysia

Roliana Ibrahim, Universiti Teknologi Malaysia

Xiaoli Li, Beijing Univ. of Technology, China

Kiwon Kwon, Korea Electronics Technology Institute, Rep. of Korea

Kwanghoon Pio Kim, Kyonggi University, Rep. of Korea

In Seop Na, Chosun University, Rep. of Korea

Jongsuk Ruth Lee, Korea Institute of Science and Technology Information, Rep. of Korea

JangHyun Kim, Sungkyunkwan University, Rep. of Korea

Organizing Committee

| Financial Chairs |

Min Hong, Soonchunhyang University, Rep. of Korea

Xibin Jia, Beijing university of technology, China

Yoosin Kim, Alticast, Rep. of Korea

| International Program Committee |

Junho Ahn, Korea National Univ. of Transportation, Rep. of Korea

Jianbin Chen, Beijing Univ. of Technology, China **Jong-Moon Chung,** Yonsei University, Rep. of Korea

Viet Sang Dinh, Hanoi Univ. of Science and Technology, Vietnam

Bobby D. Gerarado, West Visayas State University, Philippines **Rize Jin**, Tiangong University, China

Shahliza Binti Abd Halim, Universiti Teknologi Malaysia **Jongho Paik**, Seoul Women's University, Rep. of Korea

Myung-Mook Han, Gachon University, Rep. of Korea **Jun Hwang**, Seoul Women's University, Rep. of Korea

shid Khokhar, Charles Sturt University, Australia **Jeongwook Seo**, Hanshin University, Rep. of Korea

Woochun Jun, Seoul National Univ. of Education, Rep. of Korea

Massila Binti Kamalrudin, Universiti Teknikal Malaysia Melaka, Malaysia

Jung Taek Seo, Soonchunhyang University, Rep. of Korea **Namgi Kim**, Kyonggi University, Rep. of Korea

Changsoo Kim, PuKyong National University, Rep. of Korea **Jihwan Kim**, Sogang University, Rep. of Korea

Seokhoon Kim, Soonchunhyang University, Rep. of Korea **Jin Kwak,** Ajou University, Rep. of Korea

Jaewan Lee, Kunsan National University, Rep. of Korea **Ali Safaa Sadiq**, Univ. of Wolverhampton, UK

Seok-Pil Lee, Sangmyung University, Rep. of Korea **Manpyo Hong,** Ajou University, Rep. of Korea

Zulkefli bin Mansor, Universiti Kebangsaan Malaysia, Malaysia

Mazni Binti Omar, Universiti Utara Malaysia, Malaysia **Hyounkyu Song,** Sejong University, Rep. of Korea

Deepanjal Shrestha, Pokhara Univ., Nepal **In Kuk Song,** Dankuk University, Rep. of Korea

Soon Ja Yeom, Univ. of Tasmania, Australia **Wangcheol Song**, Jeju National University, Rep. of Korea

Isma Farah Siddiqui, Mehran Univ. of Engineering & Technology, Pakistan

Nawab Muhammad Faseeh Qureshi, Sungkyunkwan University, Rep. of Korea

Conference Program

Sun / Dec. 13, 2020	Room 1
16:00-18:00	VIP Reception & Greetings

Mon / Dec. 14, 2020	Room 1	Room 2	Room 3
08:00-	Registration (GF)		
08:30-11:00	Session 1	Session 2	Session 3
11:10-12:40	Opening Ceremony & Keynote Speech / Landing Ballroom A		
12:40-13:30	Lunch (12:00-13:30) / Landing Ballroom B		
14:00-16:30	Session 6	Session 7	Session 8
16:40-18:00	Session 11	Session 12	Session 13
18:30-20:00	Award Ceremony & Banquet / Landing Ballroom A & B		

Conference Program

Sun / Dec. 13, 2020	Room 1
16:00-18:00	VIP Reception & Greetings

Mon / Dec. 14, 2020	Room 4	Room 5
08:00-	Registration (GF)	
08:30-11:00	Session 4	Session 5
11:10-12:40	Opening Ceremony & Keynote Speech / Landing Ballroom A	
12:40-13:30	Lunch (12:00-13:30) / Landing Ballroom B	
14:00-16:30	Session 9	Session 10
16:40-18:00	Session 14	Session 15
18:30-20:00	Award Ceremony & Banquet / Landing Ballroom A & B	

Conference Program

Tue / Dec. 15, 2020	Room 1	Room 2
09:00-	Registration (GF)	
09:30-11:00	Session 16	Session 17
11:10-12:40	Session 18	Session 19
12:40-13:30	Lunch (12:00-13:30) / Halla Room A & B & C	
13:40-15:10	Session 20	Session 21
15:20-17:00	Session 22	Session 23

Wed / Dec. 16, 2020	Room 1
09:00-11:00	Panel Discussion: Joint Research in the Era of Pandemic

* Presentation Time

- Oral Session: 10 minutes Presentation / 5 minutes Q&A

Conference Program

Sun, Dec. 13, 2020

16:00-18:00	VIP Reception & Greetings

Mon, Dec. 14, 2020

08:30-11:00	Session 1: Smart Phone Applications and Service / Mobile Internet Computing
	Session 2: Database / Data Mining / Big Data / Mobile Object Database & Software Engineering & Architecture / Cloud Computing
	Session 3: Smart Learning / e-Learning / Learning Contents Management Systems & Future Internet & Industrial Convergence Service
	Session 4: IoT (Internet of Things) / Machine to Machine
	Session 5: Multimedia / Image Processing / HCI / Intelligent Systems

11:10-12:40	Opening Ceremony / Keynote Speech *Prof. Dr. Dayang N. A. Jawawi* */ Universiti Teknologi Malaysia, Malaysia* *Title: Computational Thinking through Educational Robotics*

12:40-13:30	Lunch (12:00 ~ 13:30)

Conference Program

Mon, Dec. 14, 2020

14:00-16:30	Session 6: Database / Data Mining / Big Data / Mobile Object Database
	Session 7: IoT (Internet of Things) / Machine to Machine & Software Engineering & Architecture / Cloud Computing
	Session 8: Wireless and Sensor Network / Security & Privacy in Internet
	Session 9: Management of Internet Application / E-Business / E-Commerce & SNS & Communications / Digital media use & effects & DLMAT workshop
	Session 10: Interactive Presentation Session Ⅰ

16:40-18:00	Session 11: Smart Phone Applications and Service / Mobile Internet Computing & Wireless and Sensor Network / Security & Privacy in Internet & Multimedia / Image Processing / HCI / Intelligent Systems
	Session 12: Multimedia / Image Processing / HCI / Intelligent Systems
	Session 13: Interactive Presentation Session Ⅱ
	Session 14: Interactive Presentation Session Ⅲ
	Session 15: Interactive Presentation Session Ⅳ

18:20-20:00	Award Ceremony / Banquet

Conference Program

Tue, Dec. 15, 2020

09:30-11:00	Session 16: KETI Workshop
	Session 17: Computational Science and Engineering Platform, and Education based on HPC
11:10-12:40	Session 18: AI and Big Data Technologies for Crime Prevention Ⅰ
	Session 19: AI and Big Data Technologies for Crime Prevention Ⅱ
12:40-13:30	Lunch (12:00 ~ 13:30)
13:40-15:10	Session 20: Interactive Presentation Session Ⅴ
	Session 21: AVIGCVPR: AI / Visual Intelligence / Big Data / Computer Vision / Pattern Recognition
15:20-17:00	Session 22: "Jeju, COVID-19, and Data Analytics"
	Session 23: Interactive Presentation Session Ⅵ

Wed, Dec. 16, 2020

10:00-11:30	Panel Discussion: Joint Research in the Era of Pandemic

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Self-Adaptive Mobile Web Service Categorization Approach for Self-Adaptive Discovery

Salisu Garba, Radziah Mohamad and Nor Azizah Saadon
School of Computing, Faculty of Engineering,
Universiti Teknologi Malaysia
[e-mail: salisu.garba@graduate.utm.my, {radziahm, azizahsaadon}@utm.my]
*Corresponding author: Salisu Garba

Abstract

This paper presents a self-adaptive mobile web service categorization approach. The proposed approach tackles the massive influx of Mobile Web Service (MWS) before discovery and selection. For this work, NLP feature-based retrieval technique is used to exploits functional semantics of the MWS for identification of the goal and tags of each MWS. Subsequently, the negative selection algorithm is introduced that uses the goals and tags features of each MWS for the categorization. The empirical evidence demonstrates that the presented solution is successful in categorizing functionally similar MWS. In addition, it could boost the performance of MWS discovery, MWS selection, and MWS-based systems reliability.

***Keywords*:** Mobile web service, Self-adaptive discovery, MWS categorization, Negative selection algorithm

1. Introduction

The current web service technologies have been extended to Mobile Web Services (MWS) due to the recent incarnation of the Internet of Things (IoT) technologies and the persistence of mobile technology. The life cycle of MWS comprises publishing, discovery, selection, composition, and invocation but the most important stage is the discovery as every subsequent stage depends on the accuracy and efficiency of the discovery [1].

MWS categorization is the use of clustering or classification methods to distributing MWS into groups such that the affinity or closeness between MWS inside one group is high and the affinity or closeness between dissimilar groups is low [2].

A number of authors have recognized the potential importance of MWS categorization in dealing with the massive influx of MWS prior to discovery so as to significantly reduces overall response time, and reduce the probability of inaccurate discovery [3], [4].

In an effort to improve the precision and lower complexity of the classical clustering algorithms, meta-heuristics approaches such as Ant-based [5], Cat Swarm Optimization [6], [7], Particle Swarm Optimization [8], Negative Selection Algorithm (NSA) [9] are widely employed.

NSA is a supervised learning algorithm that uses an iterative process to generate detectors (exemplar pattern) to models the non-self space using self samples, the detector set is used as the classification mechanism to differentiate between self and nonself [10].

A given data can be classified as normal/self data or anomalous/non-self data. The adaptation capability of NSA is obtained by the dynamic change of the detectors population. NSA needs fewer time and space resources compared to similar algorithms, moreover, prior knowledge about the data is not a prerequisite for NSA and is not constrains by large training dataset and discrete data [11]–[13].

The inconsistency of meta-heuristic approaches used in solving the MWS organization problem can be attributed to the fact that the algorithms are highly skewed towards swarm intelligence while the bio-inspired algorithms are rarely used. While the efficiency and accuracy of MWS categorization remain an issue, categorizing MWS that has little functional description increased the complexity of this problem [5]–[7].

As such, this paper presents an improved MWS organization method to boost the performance of MWS discovery and certify that the discovery mechanism has the ability to deliver the best or a near-best result in nominal computational time and not having to process the whole MWS search grid.

In this research, NLP feature-based retrieval technique is used to prepare the MWS document prior to goal extraction, the extracted goals and tags are integrated. A matrix is constructed using TF-IDF after a series of training and the multidimensional scaling (MDS) is used to reduce the dimension of the MWS document for proper similarity calculation. After that, the similarity between the reduced MWS documents is computed. Finally, similar MWS are categorized using the NSA.

The remaining of the article is prepared in the following order: The recent studies in MWS clustering approaches are deliberated in Section 2; Section 3 introduces the proposed approach, describing its two algorithms; The empirical assessment and outcomes are deliberated in Section 4; Finally, the concluding remarks are deliberated in Section 5.

2. Related Studies

A large number of existing studies in the broader literature have examined categorization approaches based on standard descriptions. Classic algorithms such as k-means [14] and agglomerative [15] are mostly adopted to categorized MWS. The authors in [16], [17] uses enhanced edge counting similarity measure and graph theory, topological metrics, pragmatics to assess service similarity and structuring the set of available services as a hierarchy of classes to deal with complications caused by the prevalence of functionally equivalent MWS which leads to enamors discovery grid.

In [18] and [19], both the description and tags of the services are extracted for training and testing. The LDA model is used to evaluate the level of similarities among MWS on the basis of the topic. The services are assigned to a single topic based on a higher probability value. Subsequently, the k-means++ (k-means with initialized centroids) is adopted to place the services in the appropriate cluster.

The author in [20] hybridized K-means & Agnes. In this method, both the services description, Web APIs, and tags extracted and combine after which the degree of similarity between the semantics of the service is computed using LDA topic model. Subsequently, k-means & Agnes (agglomerative nesting for hierarchical clustering) algorithms are hybridized to determine the categories placed the services in the appropriate cluster.

Another work by Kotekar in [7], the degree of similarity between the semantics of the service is computed using TF-IDF after which feature reduction is used to transform the MWS features accordingly. K-means & CSO algorithms are hybridized to determine the categories and placed the services in the appropriate cluster.

However, the performance of meta-heuristics clustering algorithms used in this case and other cases such as [5]–[7] suffers from the occasional inconsistency of results. though tags are considered in some cases, the difference in the number of goals in each service is not well-thought-out. Besides, dimension reduction is hardly considered despite the many and inconsistent features in the description of MWS.

From the presented review of related works, it can be seen that MWS classification is challenging, particularly for those MWS that have similar functionalities. The service

providers' diversity publishers remain briefly addressed in the literature. Moreover, the features that accurately represent the functions of MWS are hardly used to achieve accurate and efficient categorization. Therefore, a self-adaptive MWS categorization approach for self-adaptive discovery is proposed.

3. Self-Adaptive MWS Categorization Approach

The proposed self-adaptive MWS categorization approach for self-adaptive discovery is designed to boost the performance of MWS discovery and certify that the discovery mechanism has the ability to deliver the best or a near-best result in nominal computational time and not having to process the whole MWS search grid. This approach is inspired by a negative selection algorithm (NSA). The two algorithms used in this approach are discussed in the following sections.

3.1 Pre-Processing & Detector Generation

The algorithm in **Fig. 1** is used to direct the extraction of features as well as the formation of categories. Lines 4 to 8 explain how the Goals are obtained from the MWS document, coupled with Tags variables, and translated to NLTK tokens. This also explains how the category profile is constructed.

Algorithm 1: Categories formulation algorithm

1. **Input:** Crawled MWS Data/Training Set TS'
2. **Output:** Initial Detector Set DS'
3. *//Compute the features of the training sample*
4. $TS' \leftarrow \emptyset, K \geq 1$
5. **for** all Data Points in TS'
6. Pre-process using NLTK
7. Compute TF-IDF using $tf - idf_{x,S,TS} = tf_{x,S} \times idf_{x,TS}$
8. Reduce dimension using $M = [S_{ij}]$, and $S_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$
9. Compute the optimal number of K using $W_k = \sum_{r=1}^{k} \frac{1}{n_r} D_r$
10. **end for**
11. Initialize K random Centroids $C, 1 \leq i \leq k$
12. **for** all instances S in TS **do**
13. Compute the nearest Center using $Sim(S1, S2) = \frac{\sum_{i=1}^{n} S1_i \times S2_i}{\sqrt{\sum_{i=1}^{n} S1_i^2} \times \sqrt{\sum_{i=1}^{n} S2_i^2}}$
14. Assign S to K
15. Recompute Centroids C
16. **if** Centroids C changes **then** Repeat 12 and 13
17. **else** Check the Distance between S and τ in K
18. **if** $S < \tau$ **then** Remove S
19. **else** Add S to DS
20. **end for**
21. **end**

Fig. 1. Pre-Processing & formulation algorithm.

3.2 MWS Categorization

The algorithm in **Fig. 2** is used to direct the MWS categorization. The two phases of the categorization are explained in Lines 6 to 15. This also explains how the category adapts as a result of the addition or subtraction of MWS.

Algorithm 2: MWS Categorization Algorithm

1. **Input:** MWS to be classified and Input Variable $X = (x_1, x_2, \ldots x_n)$ where $\forall x \in S$
2. Initial set of detectors, N_{max} and the threshold τ
3. **Output:** Set of detectors $D = (d_1, d_2, \ldots d_n)$ that satisfy the categorization condition
4. Categories of functionally similar WS $Z = (z_1, z_2, \ldots z_n)$
5. **repeat**
6. Start Phase1: Censoring phase
7. Use the x_i generated by Algorithm 1 as initial candidate detectors d;
8. Select web service data x, where $x \in S$;
9. Calculate the similarity of x with every detector d *in* $D, \forall d \in D$ using Equation (3)
10. Start Phase2: Monitoring phase
11. Check the Distance $f_{aff}(d, x)$;
12. **if** $f_{aff}(d, x) < \tau$ **then** Add the new x to the Category Z;
13. **else** Add x *to* D;
14. **end if**
15. **until** $|D| = N_{max}$ or D cover all MWS

Fig. 2. MWS categorization algorithm

4. The Experimental Evaluation

To appraise the accuracy of the presented solution, the ProgrammableWeb dataset used in [9], [18], and [20]. The accuracy, precision, the rate of recall, and sometimes f-measure metrics are utilized to gauge the proposed MWS categorization approach. **Fig. 3** shows the empirical evaluation procedure while **Table 2** shows the evaluation metrics.

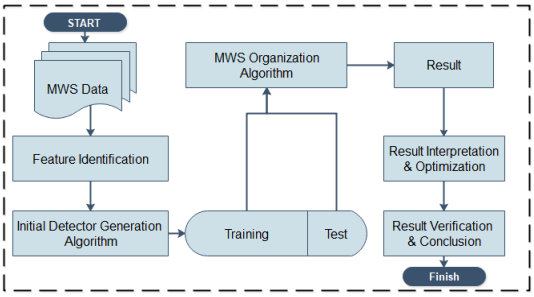


Fig. 3. The empirical evaluation procedure

Table 1. The evaluation metrics

Metric	Formula	Description
Recall	$Rec = \frac{TP}{TP + FN}$	Recall refers to the percentage of ground-truth MWS correctly categorized as members of a particular category
Precision	$Prec = \frac{TP}{TP + FP}$	Precision refers to the percentage of correctly categorized MWS as members of a particular category out of all MWS
Accuracy	$Acc = \frac{TP + TN}{TP + TN + FP + FN}$	The ratio between the number of MWS correctly categorized
F1 score	$F1 = 2 \times \frac{Pre \times Rec}{Pre + Rec}$	F1-score refers to the harmonic mean of the precision and the recall

4.1 Result and Comparison

In order to substantiate the usefulness and adequacy of the proposed solutions in tackling the problem, four state-of-the-art approaches (K-Means++ in [18], Pk-Means in [19], K-Means & Agnes [20], K-Means & CSO in [7]) are used as the benchmark approaches for comparison with the proposed approach. The specifics of these benchmark approaches are described in the related works.

Table 2. The comparison of the results

Approach	Rec	Pre	Acc	F-1
K-means++	0.470	0.470	0.529	0.470
K-means & Agnes	0.868	0.731	0.706	0.794
K-means & CSO	0.871	0.801	0.742	0.835
PK-means	0.757	0.743	0.857	0.750
K-means & NSA	**0.979**	**0.933**	**0.915**	**0.955**

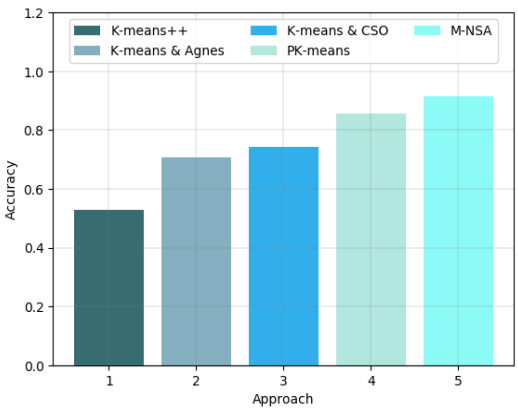


Fig. 4. The comparison of the accuracy

The comparison between baseline approaches and the proposed approach is shown in **Table 2**, while the difference in terms of accuracy is shown in **Fig. 4**. Compared with the latest baseline approaches, the proposed MWS categorization approach achieves better Accuracy of 0.915, Precision of 0.933, Recall of 0.979, and F-measure of 0.955. Through the use of a quality feature filtering method that combines MWS goals and tags leads to significant improvement in the accuracy of the categorization. Moreover, the formulation of the profile of each category provides the proposed approach with the needed intelligence to discriminate between what belongs to a particular category and what does not. This is rarely applicable to the related works.

5. Conclusions

In this paper, an improved self-adaptive MWS categorization approach for self-adaptive discovery is proposed to deal with complications caused by the prevalence of functionally equivalent MWS and boost the performance of MWS discovery. The proposed algorithms in this work are based on NSA. The algorithm is applied to the programmableWeb dataset, and the effect indicates a substantial increase in accuracy relative to contemporary studies. In the future, other approaches, such as transfer learning, can be investigated for the extraction of MWS features to improve performance prior to convergence with the self-adapted MWS discovery framework.

References

[1] H. Guermah, T. Fissaa, H. Hafiddi, and M. Nassar, "Exploiting Semantic Web Services in the Development of Context-Aware Systems," *Procedia Comput. Sci.*, vol. 127, pp. 398–407, 2018.

[2] B. Xia, Y. Fan, W. Tan, K. Huang, J. Zhang, and C. Wu, "Category-Aware API Clustering and Distributed Recommendation for Automatic Mashup Creation," *IEEE Trans. Serv. Comput.*, vol. 8, no. 5, pp. 674–687, 2015.

[3] B. Jiang, L. Ye, J. Wang, and Y. Wang, "A Semantic-Based Approach to Service Clustering from Service Documents," in *2017 IEEE International Conference on Services Computing (SCC)*, 2017, pp. 265–272.

[4] A. Bukhari and X. Liu, "A Web service search engine for large-scale Web service discovery based on the probabilistic topic modeling and clustering," *Serv. Oriented Comput. Appl.*, vol. 12, no. 2, pp. 169–182, 2018.

[5] G. Spezzano, "Using Service Clustering and Self-Adaptive MOPSO-CD for QoS-Aware Cloud Service Selection," *Procedia Comput. Sci.*, vol. 83, no. Ant, pp. 512–519, 2016.

[6] S. Kotekar and S. S. Kamath, "Enhancing service discovery using cat swarm optimisation based web service clustering," *Perspect. Sci.*, vol. 8, pp. 6–8, 2016.

[7] S. Kotekar and S. S. Kamath, "Enhancing Web Service Discovery Using Meta-heuristic CSO and PCA Based Clustering," *Prog. Intell. Comput. Tech. Theory, Pract. Appl.*, vol. 519, pp. 393–403, 2018.

[8] A. Nagy, C. Oprişa, I. Salomie, C. B. Pop, V. R. Chifu, and M. Dinşoreanu, "Particle swarm optimization for clustering semantic web services," *Proc. - 2011 10th Int. Symp. Parallel Distrib. Comput. ISPDC 2011*, pp. 170–177, 2011.

[9] S. Garba, R. Mohamad, and N. A. Saadon, "Search Space Reduction Approach for Self-adaptive Web Service Discovery in Dynamic Mobile Environment," in *Emerging Trends in Intelligent Computing and Informatics*, F. Saeed, F. Mohammed, and N. Gazem, Eds. Cham: Springer International Publishing, 2020, pp. 1111–1121.

[10] S. J. Nanda and G. Panda, "A survey on nature inspired metaheuristic algorithms for partitional clustering," *Swarm Evol. Comput.*, vol. 16, pp. 1–18, 2014.

[11] S. Fouladvand, A. Osareh, B. Shadgar, M. Pavone, and S. Sharafi, "DENSA: An effective negative selection algorithm with flexible boundaries for self-space and dynamic number of detectors," *Eng. Appl. Artif. Intell.*, vol. 62, pp. 359–372, 2017.

[12] C. Wen and L. Tao, "Parameter analysis of negative selection algorithm," *Inf. Sci. (Ny).*, vol. 420, pp. 218–234, 2017.

[13] S. Deng, L. Huang, Y. Li, and J. Yin, "Deploying Data-Intensive Service Composition with a Negative Selection Algorithm," *Int. J. Web Serv. Res.*, vol. 11, no. 1, pp. 76–93, 2014.

[14] M. Shi, J. Liu, B. Cao, Y. Wen, and X. Zhang, "A Prior Knowledge based Approach to Improving Accuracy of Web Services Clustering," in *2018 IEEE International Conference on Services Computing (SCC)*, 2018, pp. 1–8.

[15] R. A. H. M. Rupasingha, I. Paik, and B. T. G. S. Kumara, "Improving Web Service Clustering through a Novel Ontology Generation Method by Domain Specificity," in *Proceedings - 2017 IEEE 24th International Conference on Web Services, ICWS 2017*, 2017, pp. 744–751.

[16] P. Barod, M. Bhamare, and R. Patankar, "A Novel Approach for Web Service Recommendation," in *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies - ICTCS '16*, 2016, pp. 1–4.

[17] A. Corbellini, D. Godoy, C. Mateos, A. Zunino, and I. Lizarralde, "Mining social web service repositories for social relationships to aid service discovery," in *IEEE International Working Conference on Mining Software Repositories*, 2017, pp. 75–79.

[18] B. Cao *et al.*, "Mashup service clustering based on an integration of service content and network via exploiting a two-level topic model," *Proc. - 2016 IEEE Int. Conf. Web Serv. ICWS 2016*, pp. 212–219, 2016.

[19] M. Shi, J. Liu, B. Cao, Y. Wen, and X. Zhang, "A Prior Knowledge based Approach to Improving Accuracy of Web Services Clustering," *2018 IEEE Int. Conf. Serv. Comput.*, pp. 1–8, 2018.

[20] B. Cao, X. Frank Liu, J. Liu, and M. Tang, "Domain-aware Mashup service clustering based on LDA topic model from multiple data sources," *Inf. Softw. Technol.*, vol. 90, pp. 40–54, 2017.

Future Evaluation of COVID-2019 to 2021

Soobia Saeed
School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia
Skudai, Johor Bharu, 81310, Malaysia
[E-mail: soobiasaeed1@gmail.com]
*Corresponding author: saeed.soobia@graduate.utm.my

Abstract

The clinical acoustics of SARS-CoV-2 patients is critical to distinguish them from other respiratory infections. This study aims to provide an overview of the Coronavirus disease that has seriously affected the lives of people around the world. Many people have lost loved ones, and the death toll worldwide has exceeded 300 million and is still rising There is limited evidence available to demonstrate the accuracy of a CT or deep learning model that uses detection networks to detect COVID - 19 patients who are not infected with COVID - 19. Or community-acquired pneumonia. The primary purpose of this study is to present the course of the epidemic next year that will largely depend on the arrival of a vaccine and how long the immune system remains protected after vaccination or recovery from infection. Predict the future of global Coronavirus (Covid-19) cases from June 2020 to June 2021 and implement statistical equipment to measure the global normality rate. Therefore, since CT has a high diagnostic sensitivity for identifying lesions, but it is not always unique to Covid-19 and is very similar to other viral infectious diseases, it is very important to have an AI program that expresses targeted screening for lung infection on compared to affect them.

***Keywords*:** Artificial Intelligence, COVID-19, Global, Tomography

1. Introduction

University of Minnesota scientists estimate that the coronavirus outbreak will last 18 to 24 months. Researchers at the University's Center for Infectious Disease Research and Policy (CIDRAP) reported in a study published Thursday that Covid-19 was more contagious than the flu and is likely to continue to spread this spring after the first wave[1]. They said the new coronavirus has a longer time of incubation than flu, which means the symptoms would take longer to develop after a person is infected, suggesting that the coronavirus could be more infectious before symptoms appear. It was also found that , in addition to a higher R0 rating than influenza, Covid-19 displayed a greater asymptomatic transmission rate. A blank R, or R0, with a rating with 1 means a person with the virus spreads it to someone else on average[2].

The epidemic is expected to continue if no more than 1 R-value is given in the course of an event like the current pandemic. In different studies, the R0 for Covid-19 was scored between 1.4 and 5.7, but CIDRAP suggested that a score was difficult to measure due to differences in the identification and testing of infected people among regions[3].

It appears that some countries have managed to drive their R0 for (Covid-19) below 1 with mitigation measures, despite the removal of mitigation measures, the R0 could fall to above 1 in any area , resulting in a return. Diseases appear over time, the writers of the study said, adding that a higher R0 means that more people will need to be exposed and immunised until the outbreak is over. It is possible that it will not end before 60 per cent to 70 per cent of the population is resistant. The study predicts that the outbreak of Covid-19 will last about 18 months with the current flu epidemics and the pace at which the Corona virus will spread, and two years[4].

2. Three Pandemic Scenarios

Scientists have come up with three potential ways the virus could continue to spread in the years to come. In the first example, the first Coronavirus wave will persist until spring 2020

and will be followed by a "series of smaller regular waves" that occur over a span of one to two years during the summer and then continuously. The virus would then eventually subside sometime in 2021[5].

Analysts said the occurrence of these waves may vary globally, depending on the mitigation measures being applied and how they can be mitigated. This scenario may involve periodical renewal and subsequent relaxation of mitigation measures over the next year or two, depending on the height of the wave crests[6-7].

The second example would see the Coronavirus' first wave in the spring followed by a greater wave in the fall or winter of 2020. According to the study one or two smaller waves will arrive later in 2021. The researchers noted that this trend was close to that seen during the epidemic of flu in 1918, which killed 50 million people worldwide[8-9]. A first "burning" wave will proceed in the above scenario for an in progress transmission but without a specific pattern of wave. The pattern may be geographically different and may be influenced by the degree of mitigation measures in different regions. The authors of the study, however, noted that this condition is unlikely to require the reintroduction of prevention steps, while new cases and deaths have persisted[10]. If the pandemic scenario follows, we must be prepared for at least another 18 to 24 months of prolonged Covid-19 activity, with hotspots frequently emerging in different areas, assuming at least some degree of mitigation measures are underway. They added that the Coronavirus is likely to continue circulating among humans even as the epidemic recedes. As its intensity decreases over time , the study predicted that "the virus coincides with a seasonal trend". Research is in full swing across the globe on developing a vaccine or treatment for Covid-19[11]. According to the World Health Organisation, there are currently at least 102 vaccines under development worldwide, but experts have anticipated the vaccine to take 12 to 18 months to enter the market. Meanwhile Gilead Sciences claims to have seen "strong evidence" as a treatment for Covid-19 from clinical trials of its drug remdesivir. The medication has not been officially approved to treat the virus, however, and US health officials alert that the results have not yet been peer reviewed[12].

3. Prediction Cases and Death

Coronavirus disease has had a significant effect on people's lives around the world. So many people have lost loved ones, and the death toll worldwide has exceeded 47,000 and is still increasing. 699,000 recent casualties were publicly announced and confirmed in 2020. The outlook for coronavirus cases in 2021 is declining in the future, but we now hope that attempts to avoid a second major epidemic, such as social distancing, will continue as long as possible.

4. Comparative Evaluation

The evaluation criteria will be learned and explained, and the findings and outcomes will then be presented in the current methods after applying the evaluation criteria. The purpose of this analysis is to provide a ranking and comparison with the design of the most prominent research methods to verify the ability of the predominant test approaches to improve the test of statistical regression.

4.1 Evaluation Results

This section explains the findings of the first scientific statistical methods that also verified that our version became successful and robust in following patients who were seriously infected with COVID-19. With this experience, we explore in-depth knowledge of automated tomography (CT) scans to turn infectious COVID-19 areas into pulmonary fieldsSecond, compared to detection by handwriting, this approach offers quick and reliable separation of COVID-19 contagious areas. Our devices provide a comprehensive approach to monitoring disease progression and analysing improvements in the period of COVID-19 to the course of care with the necessary quantitative records. We are therefore optimistic that this deep master machine measuring COVID-19 to 2021 can talk with confidence to this network with several new studies on interesting features. This prospective study focuses on retrospective progress in the follow-up of COVID-19 patient CT tests and evaluations. In the quantification of a particular COVID-19 related insect.

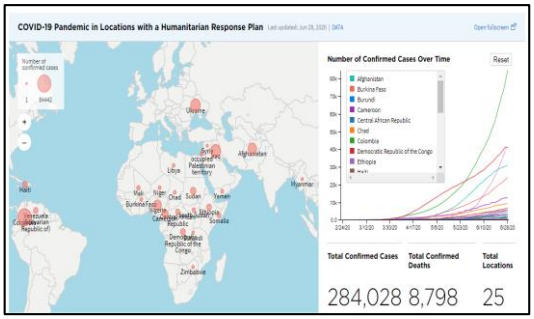


Fig. 1. This image shows the number of reported cases and deaths from Coronavirus (COVID-19) on the Human Response Plans (HRP) websites. COVID-19 is information from the World Health Organization (WHO).

5. Conclusions

The goal of this research is to provide an overview of the Coronavirus disease that has seriously affected people's lives around the globe. Lots of people have lost loved ones, the death toll around the world exceeds 300 million and continues to rise. There is insufficient evidence available to demonstrate the accuracy of a CT scan or deep learning model utilising networks of detection to detect / detect COVID-19 patients among those without COVID-19 or community-acquired pneumonia. The direction of the 20201 outbreak next year will rely largely on the arrival of a vaccine and how long the immune system remains healthy from infection after vaccination or recovery. For decades, many vaccines, such as those against measles or polio, have been protective although disappearing over time, such as whooping cough and influenza. Long-term immunity is also produced by some forms of viral infections, and others by a transient response. This period of immunity will critically depend on the overall incidence of SARS-CoV-2 through 2025.

References

[1] Badr, Hamada S., Hongru Du, Maximilian Marshall, Ensheng Dong, Marietta M. Squire, and Lauren M. Gardner. "Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study." *The Lancet Infectious Diseases* (2020).

[2] Silal, Sheetal, Juliet Pulliam, Gesine Meyer-Rath, Brooke Nichols, Lise Jamieson, Zaid Kimmie, and Harry Moultrie. "Estimating cases for COVID-19 in South Africa Update: 19 May 2020." *Update* (2020).

[3] Nouvellet, P. et al. Report 26: Reduction in Mobility and COVID-19 Transmission https://doi.org/10.25561/79643 (Imperial College London, 2020).

[4] Zhou, Ying, Renzhe Xu, Dongsheng Hu, Yang Yue, Qingquan Li, and Jizhe Xia. "Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: a modelling study using mobile phone data." *The Lancet Digital Health* 2, no. 8 (2020): e417-e424.

[5] Santana-Cibrian, Mario, Manuel A. Acuna-Zegarra, and Jorge X. Velasco-Hernandez. "Lifting mobility restrictions and the induced short-term dynamics of COVID-19." *medRxiv* (2020).

[6] Moriyama, Miyu, Walter J. Hugentobler, and Akiko Iwasaki. "Seasonality of respiratory viral infections." *Annual review of virology* 7 (2020).

[7] Flaxman, Seth, Swapnil Mishra, Axel Gandy, H. Juliette T. Unwin, Thomas A. Mellan, Helen Coupland, Charles Whittaker et al. "Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe." *Nature* 584, no. 7820 (2020): 257-261.

[8] Havers, Fiona P., Carrie Reed, Travis Lim, Joel M. Montgomery, John D. Klena, Aron J. Hall, Alicia M. Fry et al. "Seroprevalence of antibodies to SARS-CoV-2 in 10 sites in the United States, March 23-May 12, 2020." *JAMA Internal Medicine* (2020).

[9] Zhao, Juanjuan, Quan Yuan, Haiyan Wang, Wei Liu, Xuejiao Liao, Yingying Su, Xin Wang et al. "Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019." *Clinical Infectious Diseases* (2020).

[10] Moore, Kristine A., M. Lipsitch, John M. Barry, and M. T. Osterholm. "COVID- 19: The CIDRAP Viewpoint: Part 1: The Future of the COVID- 19 Pandemic: Lessons Learned from Pandemic Influenza." *CIDRAP: University of Minnesota. April 30th* (2020).

[11] Davies, Nicholas G., Petra Klepac, Yang Liu, Kiesha Prem, Mark Jit, Rosalind M. Eggo, and CMMID COVID-19 working group. "Age-dependent effects in the transmission and control of COVID-19 epidemics." *MedRxiv* (2020).

[12] Long, Quan-Xin, Bai-Zhong Liu, Hai-Jun Deng, Gui-Cheng Wu, Kun Deng, Yao-Kai Chen, Pu Liao et al. "Antibody responses to SARS-CoV-2 in patients with COVID-19." *Nature medicine* (2020): 1-4.

SECURING INTERNET BANKING TRANSMISSION IN THE ERA OF HIGH-SPEED COMPUTING: A CRITICAL REVIEW FOR BETTER CONFIDENTIALITY AND NON-REPUDIATION

Muhammad Saidu Aliero[1], Imran Ghani[2], Muhammad Fermi Pasha[3]
[1,3]School of Information Technology, Monash University Malaysia
[2]Indiana University of Pennsylvania, Philadelphia, USA

Abstract

Internet banking is a system which allows bank clients to access their accounts and its services through a laptop or smart phone or any other devices. Nowadays, customers more and more request and use internet banking over internet to make financial transactions. Indeed, Internet banking has a lot of advantages such as the speed and flexibility but however there are a lot of security challenges. The transmission during internet banking should be secure from potential attackers. Internet is a public network and open which makes it an insecure channel. Encryption is the procedure of scrambling data into cipher format which can provide more protection to transmitted data over unsecured channels such as the Internet. Cryptography is used to provide encryption. The current cryptography algorithms are used to achieve the confidentiality, non-repudiation and some other requirements. In this paper, the existing cryptography and its methods and objectives have been discussed. These existing methods used to secure the transmission of internet banking have been critically reviewed in the literature. A critical review of related works reveals that available solutions are poor in terms of achieving confidentiality and non-repudiation in this era of High-Speed Computing. Furthermore, a security solution has been recommended to overcome the mentioned problems.

***Keywords*:** Internet banking; cryptography; confidentiality; non-repudiation

1. INTRODUCTION

Internet banking or online banking can be defined as "the provision of financial services and markets using electronic communication and computation" [1],. Nowadays, online Banking or internet banking has become a very important feature of modern banking. In this age of technology, information and communications technology (ICT) is playing an important role in the business [2-4]. Bank customers are demanding and requesting more and more services through online banking. Banks are highly competing to provide attractive services online. So, they distribute their services through insecure channels such as Internet. Actually, there are a lot of vulnerabilities and potential attackers during the transmission over internet. Cryptography and specifically encryption methods is a way of getting a secure transmission over unsecured channels such as the Internet [5]. The current cryptography algorithms are used to achieve the confidentiality, non-repudiation and some other requirements.

Internet banking and its history are discussed in this paper. The cryptography methods can be classified in to symmetric and asymmetric methods. These methods have been clarified in this work. Then, confidentiality and non-repudiation have been explained. This article sheds light on existing security solutions of internet banking during the transmission. **Fig. 1** shows the organization of this critical review paper.

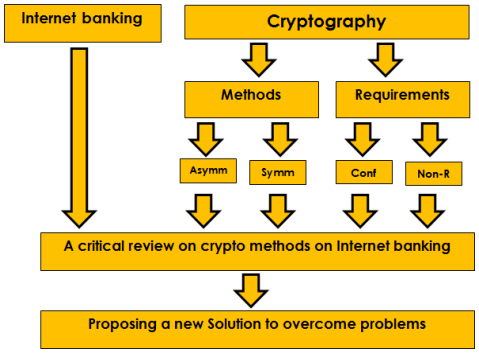


Fig. 1. organization of research

In this paper, the existing cryptography and its methods and objectives have been discussed. These existing methods used to secure the transmission of internet banking have been critically reviewed in the literature. A critical review of related works reveals that available solutions are poor in terms of achieving confidentiality and non-repudiation in the era of High- Speed Computing. Furthermore, a security solution has been recommended to overcome the mentioned problems.

2. INTERNET BANKING

Online banking system became popular in 1980s in New York, when distance service was provided for the purpose of financial transaction through phone line connected to keyboard and TV monitor. It allows on- line viewing of statements, bank transfers and bill payments. In order to make bank transfers and bill payments, a written instruction giving details of the intended recipient had to be sent to the NBS who set the details up on the Home link system. This service allows customer to perform financial transaction in affordable price. However, it was partially accepted by people when clicks and bricks euphoria hit in the late 1990. This is because of the fair loyalty from both customer and service provider for dealing with financial transactions online despite the benefit of online banking transaction compared to traditional banking system.

In 2001, most organization such as eBay, Amazon America Online adopted web-based applications as backbone for conducting their business in which 80% of US bank offered online banking. Bank of America became the first bank to top 3 million online banking customers, more than 20 percent of its customer base. In 2009 Gartner group published a report saying that approximately 47% of adults in U.S and 30% of adults in U.K are using online banking system [6, 7].

CRYPTOGRAPHY

Cryptography is the practice and study of techniques that enables end parties to communicate securely with the presence of third parties, ensuring that third parties does not deduce communication mechanism used and information that is being transferred between the end parities. It is more specifically on constructing and analyzing protocols that prevent third parties from breaching the goal of information security, such as confidentiality, integrity, availability and non-repudiation which are the central focus of modern cryptography. Modern cryptograph is a combination of elements from the field of mathematics, computer science and electrical engineering. The application of cryptography includes ATM cards, computer passwords, and electronic commerce.

Before the advent of modern age, cryptography was interchanged with encryption. That is transformation of text from readable form into unreadable form. The encoder of the message needs to share decoding techniques to intended parties to prevent third parties from decoding the message. This method of cryptology was stared in World War I and became increasingly complex and widespread in its application with the advent of computer system [8].

2.1 CONFIDENTIALITY

Confidentiality is one of the key concerns in information security environment. It is characteristic of information that needs to be kept secret from unauthorized access. Confidentiality of information is breached when company gain access to the other company"s secret recipe. Secret formulae, military intelligence data and bank information need to be kept secret from unauthorized access. The goal of confidentiality is to make sure information system and information are kept secret from unauthorized accessing, copying and disclosing information.

Protection of confidentiality is not limited between company-to-company or government-to-government rather, it covered all kinds of information that are not intended to be seen by third parties. Confidentiality can be achieved using one of the following cryptographic algorithms such as, Data Encryption Standard (DES), Triple Data Encryption Standard (3DES), Advanced Encryption Standard (AES), Rivest, Shamir and Adleman cryptography algorithm (RSA), Elliptic curve cryptography (ECC), NTRU cryptography algorithm (NTRU) or any other cryptography algorithm [8,9].

2.2 NON-REPUDIATION

In the field of information security, the sender of information cannot deny sending of information and so also the receiver of the information cannot deny receiving it. Most financial institutions use technology such as digital signature and encryption to achieve the user's authenticity and ascertain non-repudiation. With the non-repudiation factor, one cannot deny involvement. Non-repudiation refers to the obligation to agree that they are involved in exchanging the data [8].

2.3 CRYPTOGRAPHY METHODS

Cryptograph theoretically is supposed to provide 100% secrecy to information, but practically is not. Since the First World War I, there had been extensive research effort made in the field of information security, and various techniques were used in many applications. Now in most electronic communications, the cryptographic technique or combination are used to achieve secure communication. These techniques can be categorized into asymmetric and symmetric cryptography. Each technique has advantages and disadvantages but however, the strength of each technique depends on how the secret key is chosen and managed between the parties. When the sender of the information encrypts a message and sends it to the receiver, the receiver has to use the same or opposite key that was used during encryption to make the content of the message clear [8].

2.4 SYMMETRIC METHODS

This is a class of cryptograph that uses one secret key for encryption and decryption and this means that the key used for encryption is the same as the key that will be used for decryption. When the sender encrypts the message using the symmetric approach, he usually includes the secret key that the receiver will use to decrypt the message (chipper text). Asymmetric cryptography encrypts data in blocks, which makes it faster than an asymmetric algorithm. However, the symmetric algorithm cannot achieve some information services such as digital signature and secret key exchange. Example of such an algorithm is DES, 3DES, and AES [8].

2.5 ASYMMETRIC METHODS

This is a class of cryptography that uses different encryption and decryption. In this case, each member of the party will have two different keys i.e. one for encryption and one for decryption, known as the private and public key. Usually, when user registers to use encryption, two keys are present to him and to

be able to communicate with the other member of the party. He needs to share his public key so that they would be able to send him an encrypted message. The sender usually encrypts messages to be sent by the receiver public key and the receiver uses his private key to decrypt the encrypted message sent by the sender. Unlike in asymmetric cryptograph, the same key is used for both encryption and decryption purposes. Now in modern technology symmetric key are used to implement security in network-related application than using asymmetric due to it fastness and easy to manage the key. However, the advent of symmetric does not eliminate the use of asymmetric because it is used to verify the original sender's integrity and share symmetric keys through the networks [8].

3. A CRITICAL REVIEW ON EXISTING SOLUTIONS

For a clear and comprehensible vision, the review of previous works has been tabulated as shown in **Table 1**.

As it has been mentioned above, there are some researchers who proposed security techniques to secure the internet banking transmission, and they suggested RSA as a digital signature method such as [11, 14, 16, 18]. While some other cases, researchers had used ECC crypto method as mentioned in [10 15, 19] However, more recent studies have shown that RSA and ECC cannot resist the high-speed computing [20-23]. In addition, there are some who had used AES to obtain confidentiality such as [12, 13]. However, Non-repudiation would not be achieved by using any of symmetric methods. In a nutshell, there is a need to improve and create a new security solution to achieve the confidentiality as well as Non-repudiation in this era of quantum computing.

4. PROPOSING AND CONCLUSION

In conclusion, the ISO/IEC JTC1/SC27 standard stated that hybrid cryptography can be defined as the branch of asymmetric cryptography that uses suitable symmetric techniques to reduce some of the problems inherent to normal asymmetric cryptosystems (e.g., the problems met once trying to transmit long messages rapidly). The hybrid cryptography also can be illustrated as the area of asymmetric cryptography which applies keyed symmetric crypto methods as black-box methods with specific security requirements [23]. Contrasting symmetric methods, asymmetric methods would be able to provide integrity, authentication and non- repudiation as well as confidentiality. It is trustworthy to mention that symmetric methods are faster than asymmetric methods. For some problems that need just confidentiality, symmetric algorithms are the best option. However, for problems related to other security requirements such as integrity, authentication and non-repudiation, one of the asymmetric methods are required to be used. Therefore, hybrid crypto methods would help to obtain the security requirements of symmetric side by side with asymmetric methods.

In this paper, two main targets are considered to improve the security of internet banking. The first objective is to ensure a high level of secrecy and confidentiality of internet banking in the era of high speed computing. Non-repudiation is the second objective of this research. A high speed computing computer or a quantum computing is a device for computation that makes direct use of quantum mechanical phenomena to obtain operations on information. High speed computers are different from normal computers based on transistors. The idea of high speed computing is that quantum properties can be used to represent data and perform processes on this information. AES crypto method and NTRU crypto method are considered stronger in Quantum Resistance than other symmetric and asymmetric methods. Thus, the authors strongly recommend that using a hybrid of AES and NTRU would help to improve the security transmission of internet banking in the era of high speed computing which would be able to achieve efficiency, confidentiality and non-repudiation.

Table 1. A critical review on existing solutions for secure transmission of internet banking

Author (s)	Contribution	Cryptography Methods	Security Objective
[10]	They develop an encryption based method for data confidentiality. In addition, Digital signature for data authentication and non-repudiation has been developed.	AES, ECC	Confidentiality Non-repudiation
[11]	They developed a method by using User anonymity and location un-traceability through a series of anonymizing host or mixes. Payment transaction un-traceability using a blind signature based on RSA encryption method. Confidentiality of payment transaction using dual and digital signature	RSA	confidentiality
[12]	They provide an authentication protocol which consists of 3 main phases: registration, login and recovery. The registration phases allows a user and the bank to generate a shared secret key which will be used to encrypt the registration of SMS. In the login phase, the user produces a one- time through his mobile phone and delivers other information in an encrypted format to the Bank Server via SMS. In the recovery phase, a process to retrieve his/her Opass setting is enable in case of loss of a mobile phone	AES	Confidentiality
[13]	They developed a combination of a Biometric technique and cryptographic algorithm to increase the security of the transaction. Apart from a PIN, a One Time Password is sent to the user"s mobile phone to verify the identity of the user. To help prevent any eavesdropping, a stenography method used a fingerprint image is applied, which can be prevent a man in the middle attack..	AES	Confidentiality
[14]	They develop a system that assures the authentication of the customer with the help of a dual secure system composed of a PIN number and ATM card. They used an asymmetric encryption to secure transaction between the mobile"s user and the Bank Server.	RSA	Confidentiality

[15]	They developed a system that enable user to apply for an E-check document from the Bank which can achieve Confidentiality, and Non-repudiation. Once this E-check is issued to the user, the user will use it to process a secure transaction from his account to the merchant"s account. E-check is based on the user's account in the bank and the user"s account is signed with user's private key to ensure authentication and non-repudiation. Each E-check is signed by bank with private key to ensure that it cannot be forged.	ECC	confidentiality, and Non- repudiation
[16]	The system uses a tamper-resistant crypto device as client authentication factor to ensure the security of Internet Banking transactions. If the user wants to login Internet Banking account, the Token will verify the user"s PIN at first, and will sign the transaction data with the RSA private key stored in the Token to ensure the security and non repudiation of the transaction.	RSA	Non- repudiation
[17]	The authors use ECC algorithm combines with a biometric technology for secure One Time Password solution to achieve Confidentiality and Non-repudiation.	ECC	Confidentiality and Non- repudiation
[18]	They develop a system using AES and RSA to achieve Non- repudiation, Confidentiality security requirements. The registration process includes filling a form at a Bank desk, then receiving (after almost one week) an email from the bank to activate the access to the application. After activation, another email is sent, providing a link that must be followed in order to set the initial password. After setting the initial password, the user can login using his/her credentials.	AES RSA	Non- repudiation, Confidentiality
[19]	The author uses ECC method to obtain the confidentiality without involving the digital signature. In addition, he also uses the authenticated encryption scheme to design an electronic payment system.	ECC	confidentiality

References

[1] Allen, F. James M. and Philip S. 2002 E-finance: an introduction. Journal of Financial Services Research 22. (1) : 5-27.

[2] 2 Aziz M. and Nooh, M. 2013. Design Analysis of CIMB Bank‟s Website. Sains Humanika, 66(1).

[3] 3 Al-Ghazali, B, Rasli, A and Yusoff,. 2015. A Review of Theories on Technology Acceptance: The Case of Mobile Banking User Retention in Saudi Arabia. Jurnal Teknologi, 72(5).

[4] Amiri, I. S., Afroozeh, A., Bahadoran, M., Ali, J., & Yupapin, P. P. (2012). Molecular transporter system for qubits generation. Jurnal Teknologi, 55(1), 155-165.

[5] Salleh, M. Ibrahim, S. and Isnin, I. F. 2012. Image encryption algorithm based on chaotic mapping. Jurnal Teknologi, 39(1), 1-12.

[6] Cronin, Mary J., ed. Banking and Finance on the Internet. John Wiley & Sons, 1997.

[7] Keivani, F. Sameni, M. Jouzbarkand, M. Khodadadi, and Z. Khalili Sourkouhi. 21012. A General View on the E-banking. International Proceedings of Economics Development & Research.

[8] Menezes, Alfred J., Paul C. Van Oorschot, and Scott A. Vanstone. Handbook of applied cryptography. CRC press, 1996.

[9] Sidek, M., Rahim, A., & Sha'ameri, A. Z. (2007). Comparison analysis of stream cipher algorithms for digital communication. Jurnal Teknologi, 46(D), 1-16.

[10] Jin, N. & Xianling, H., 2008. Mobile banking information security and protection methods. Proceedings - International Conference on Computer Science and Software Engineering, CSSE 2008, 3, pp.587–590.

[11] Barskar, R. et al., 2010. The Algorithm Analysis of E- Commerce Security Issues for Online Payment Transaction System in Banking Technology. , 8(1), pp.307–312. Available at: http://arxiv.org/abs/1005.4266.

[12] Joseph, M., Mathew, J.J. & Mathews, D., 2014. Securing Internet Banking Authentication Through oPass. , pp.78–81.

[13] Murugesh, R., 2012. Advanced biometric ATM machine with AES 256 and steganography implementation. 4th International Conference on Advanced Computing, ICoAC 2012, pp.0–3.

[14] Rajanbabu, D. and Raj, C. 2014. Implementing a reliable cryptography based security tool for communication networks. International Conference on Science, Engineering and Management Research (ICSEMR 2014). ISBN 9781479976133. IEEE-32331.

[15] Xierui, 2010. Secure E-check payment model based on ECC. Proceedings - 2010 WASE International Conference on Information Engineering, ICIE 2010, 2, pp.109–112.

[16] Zhou, Y. and Guo, S., 2011. SPA-based security evaluation of RSA implementation in Internet banking USB token. Proceedings - 2011 International Conference on Instrumentation, Measurement Computer, Communication and Control, IMCCC 2011, pp.499–503.

[17] Mahto, D., 2015. Enhancing Security of One-Time Password using Elliptic Curve Cryptography with Biometrics for E-Commerce Applications. , pp.1–6.

[18] Itan, V.G.A. and Lupu, V., 2015. Security enhancement of internet banking applications by using multimodal biometrics. , pp.47–52.

[19] Yang, J., Chang, Y. & Chen, Y., 2013. An Efficient Authenticated Encryption Scheme Based on ECC and its Application for Electronic Payment. , 42(4), pp.315–324.

[20] 5. Jarvis K, NTRU over the Eisenstein Integers. 2011. Carleton University. URL: http://www.ruor.uottawa.ca/en/handle/10393/19862 [accessed 2012-04-24] [WebCite Cache]

[21] 6. Stebila D, Mosca M, Lütkenhaus N. The case for quantum key distribution. Quantum Communication and Quantum Networking, 2010: 283-296. URL: http://arxiv.org/abs/0902.2839 [accessed 2012-04-25][WebCite Cache]

[22] Alanazi, H, Jalab, H, Alam, G. Zaidan B. Zaidan A . 2010. Securing electronic medical records transmissions over unsecured communications: An overview for better medical governance. Journal of Medicinal Plants Research, 4(19), 2059-2074.

[23] Alanazi, H. O., Zaidan, A. A., Zaidan, B. B., Kiah, M. M., & Al- Bakri, S. H. (2015). Meeting the Security Requirements of Electronic Medical Records in the ERA of High-Speed Computing. Journal of medical systems, 39(1), 1-13.

KSII The 12th International Conference on Internet (ICONI) 2020.

Personal Authentication using Corresponding Relationship between Objects

Gi-Chul Yang
Department of Convergence Software, Mokpo National University
Mokpo, South Korea
[e-mail: gcyang@mokpo.ac.kr]
*Corresponding author: Gi-Chul Yang

Abstract

The graphical password was first proposed by G. Blonder, and images are recalled better than words. This paper introduces and describes an efficient personal authentication scheme based on graphical passwords. The proposed system, called 'ACRO' (Authentication using a Correspondence Relationship between Objects), establishes a bulwark against shoulder-surfing attacks and is easy to use. ACRO would work well in a VR context where text based passwords are not very practical or secure and ACRO may overcome those issues.

***Keywords*:** Personal Authentication, Security, Graphical Password

1. Introduction

The graphical password was first proposed by G. Blonder [1], and images are recalled better than words [2,3]. These days, there are diverse schemes for graphical passwords [4-9]. ACRO (**A**uthentication using **C**orresponding **R**elationship between **O**bjects) introduces a new method of secure and efficient graphical password authentication. The important features of ACRO are the use of a correspondence relationship between objects and the transparent image moving technique. In addition, ACRO utilizes color information to increase usability. Color, however, is not the only criterion or indicator for identifying object relations. Color is only an auxiliary indicator to increase the usability.

ACRO introduces a scheme that can lead to the development of the ultimate graphical password system. ACRO increases usability while maintaining strong security. It increases usability by facilitating quick logins: a user only needs to move a password object (i.e., a number) in two-dimensional space. In addition, the numbers in the login interface will always appear in ascending order to allow a user to locate and predict the position of the password number. ACRO maintains strong security by using a correspondence relationship between objects within the system.

Fig. 1 shows the registration (a) and login (b) interfaces of ACRO. ACRO uses three types of objects: symbols, letters, and numbers. (The following symbols are used in the example:  .) Ten numbers (zero through nine) are used in the ACRO interface, and five numbers will appear in each row in the interface. The numbers will always be arranged in descending order from left to right, top to bottom. (The sequence will repeat like a circular buffer.) This will help the user to place the number easily in the correct position. However, the numbers do not appear at the same location

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2017R1D1A1B04032968).

for each authentication attempt. This randomness increases security power.

(a) (b)

Fig. 1. Interfaces of ACRO

Authentication proceeds in two steps: the password registration step and the password authentication step. At the password registration step, a user chooses one target image (e.g.,) and numbers (as many numbers as a user wants, for example 5, 2, and 7 in sequence) and one English letter (e.g., C). A user needs to memorize those three types of objects. A user can create his/her password by using the spatial relationship among these objects. That means the chosen number (e.g., 5) should be moved within the grid in the login interface in order to be considered a password number.

At the password authentication step, a user should move the first chosen number (i.e., 5) which appears on the same row with English alphabet C to under the target image by using left and/or right direction key(s). If a user presses the direction key, all numbers displayed on the interface move together, not just the numbers on the same row as the chosen letter. This prevents an attacker from discovering the user's password number by distributing attacker's attention.

Once the number is located at the correct symbol position, the user presses the OK button to complete the input of the first password number. This process is subsequently repeated for each password number. After inputting all of the password numbers, the user presses ENTER to log in. Then, ACRO compares the user's completed input to the registered password and processes the authentication.

2. An Example of Password Registration and Login

This section gives an example of password registration and login. The registration interface is shown in Fig. 1 (a) above. In the example interface, the symbols are , the letters range from A to E, and the numbers range from zero to nine. First, the user enters an ID. At the beginning of the password registration, a user chooses a symbol, a letter, and as many numbers as the user desires. In this case, the user chooses the symbol , the letter C, and the numbers 2, 5, and 3 as his/her password. The user presses , C, and 2, 5, 3. After selecting all the numbers, the user presses DONE to finish the registration process.

The login interface is shown in Fig. 1 (b) above. To login, the user first places the number 2 in row C under the symbol by using the left and right direction keys. Then, the user presses OK button. The user repeats this process two more times for the numbers 5 and 3. After the last number, 3, is entered, the user clicks the ENTER button.

ACRO compares the information acquired from the user with the registered information. In this example case, the entered password matches the registered password, the user is authenticated, and the login is completed.

3. Password Space of ACRO

The password space of ACRO is calculated as: The number of possible passwords, out of 25 numbers (presented in the interface shown in **Fig. 1**) taken one number at a time with three numbers, as a password will be:

$$25 * 25 * 25 = 15{,}625 \quad (1)$$

However, the above figure (1) is without considering 5 target images, 5 alphabets and 10 numbers from zero to nine for each row. Considering target 5 images, 5 alphabets and 10 numbers causes ACRO's passwords (AP) possible to be:

$$AP = (10 * 5 * 5) * (10 * 5 * 5) * (10 * 5 * 5) = 15{,}625{,}000 \text{ Passwords.} \quad (2)$$

The expression (2) shows the password space of ACRO demo system described in previous sub section. For ACRO the general password space is calculated as:

The X × Y is the size of the grid with numbers and N is the number of password digits selected. Therefore, the password space P for ACRO is:

$$AP = (10 * X * Y)^{N} \quad (3)$$

The number 10 is there in (3) since ACRO use 10 numbers from 0 to 9. For example, consider the size of the grid is 7 × 7; the password consists with eight numbers then

$$AP = (10 * 7 * 7)^{8} = 1.91581231e+20 \quad (4)$$

ACRO has a very large password space as shown in (4), so ACRO provides a strong security against brute force attack. It is observed that ACRO provides strong security against shoulder-surfing attack by moving the image transparently. At every login the position of numbers will vary. Thus it makes more difficult for any person to guess or crack ACRO password by observing. Also, randomness of object (i.e., number in each grid for the demo system) placing in the interface grid makes the attacker can get confused if he/she is trying to memorize the password details. In this way ACRO is strongly resistant to shoulder-surfing attack, while takes shorter login time than alternatives.

4. Conclusions

Graphical passwords are an alternative to current alphanumeric passwords. However, they are vulnerable to shoulder-surfing attacks. ACRO utilizes both correspondence relationship between objects and the transparent image moving technique to create a more usable and secure authentication system than other pre-existing graphical password systems. Correspondence relationship between objects makes ACRO more secure by hiding the structure of user's credentials from others. Also, it makes the attacker can get confused. Overall ACRO system is secure and resistant to various attacks.

ACRO provides high usability along with quick login time. ACRO use numbers appearing in ascending order in the interface to help a user to find a password number easily. In addition, ACRO use only two direction keys, and it reduces a login time. Hence, we can say that ACRO leads one step closer to the ideal graphical password system by increasing usability while maintaining a strong security.

References

[1] G. Blonder, "Graphical passwords," in Lucent Technologies, Inc., Murray Hill, NJ, U. S. Patent, Ed. United States, 1996.

[2] R. N. Shepard, "Recognition memory for words, sentences, and pictures," Journal of Verbal Learning and Verbal Behavior, vol. 6, pp. 156-163, 1967.

[3] A. Paivio, Rogers, T.B., Smythe, P.C., Why are pictures easier to recall then words? Psychonomic Science 11 (4), 137–138, 1976.

[4] R. Dhamija and A. Perrig, "Deja Vu: A User Study Using Images for Authentication," in Proceedings of 9th USENIX Security Symposium, 2000.

[5] S. Akula and V. Devisetty, "Image Based Registration and Authentication System," in Proceedings of Midwest Instruction and Computing Symposium, 2004.

[6] D. Weinshall and S. Kirkpatrick, "Passwords You'll Never Forget, but Can't Recall," in Proceedings of Conference on Human Factors in Computing Systems (CHI). Vienna, Austria: ACM, 2004, pp. 1399-1402.

[7] W. Jansen, "Authenticating Mobile Device Users Through Image Selection," in Data Security, 2004.

[8] S. Wiedenbeck, J. Waters, J. C. Birget, A. Brodskiy, and N. Memon, "PassPoints: Design and longitudinal evaluation of a graphical password system," International Journal of Human Computer Studies, 63, pp. 102-127, 2005

[9] G. Yang, "Development Status and Prospects of Graphical Password Authentication System in Korea," KSII Transactions on Internet and Information Systems, vol. 13, no. 11, pp. 5755-5772, 2019. DOI: 10.3837/tiis.2019.11.026.

Eliciting User Requirements for Designing Mobile Application of Eco-Care Drainage with Waste Trap Device

Nor Azman Ismail[1], Shukur Abu Hassan[2], Mohd Sharul Hafiz Razak[3], Aimi Rusdi[4], Nurul Najiha Hamdan[5] & Aqilah Hanim Mohd Taufik[6]
[1,3,4,5,6]School of Computing, Universiti Teknologi Malaysia (UTM), Johor, Malaysia
[2]School of Mechanical Engineering, Universiti Teknologi Malaysia (UTM), Johor, Malaysia
[e-mail: azman@utm.my, shukur@utm.my, sharulhafiz@utm.my, aimirusdi@gmail.com, najihah1414@gmail.com, qylahanym98@gmail.com]
*Corresponding author: Nor Azman Ismail

Abstract

In this paper, we explore users' insights that contribute to the occurrence of floods caused by clogged drains and identify user requirements in designing an eco-care smart drainage mobile application. Focus Group Discussions (FGD) was conducted with ten participants. Open-ended questions were created and asked to stimulate the discussions more strategically. Transcripts of each discussion session were recorded, categorised, and analysed thoroughly based on six themes. Based on focus group discussion and survey, results indicated that flood caused by clogged drains could be prevented through a mobile application system and waste trap device. The possible system features and components of the related application have been highlighted.

***Keywords*:** Waste Trap, Mobile App, Water Quality, Solid Waste, Suspended Solids

1. Introduction

Solid waste and gross pollutants such as litter and debris are unattractive by nature. Their presence in drains and channel flows causes many problems such as damage to natural habitat, degradation of water quality, an increase of flash flood problems, being aesthetically unpleasing and reducing amenity values [1]. Solid waste production is not a new phenomenon but has existed for centuries and continues to this day. Fortunately, solid waste management was introduced in the late 18th century [2]. Solid waste contains biodegradable elements and toxic, hazardous components which has a high cost for proper disposal [3]. Significant effects that can be seen as a result of unplanned solid waste management is poor water quality as solid waste has entered drainage [4].

To reduce the impact of solid waste pollution on drainage, innovative solid waste traps were introduced [5]–[7]. This waste trap is a device that avoids solid waste from entering water bodies which will improve not only the water quality but also the environment which traps up to 255kg rubbish per day during monsoon period [6]. More acceptable pollutants such as dirt, chemicals, heavy metals and bacteria will have the data sent to the authorities detected by a chemical sensor. Gross Pollutant Traps in the form of waste traps are one of the suggested structural Best Management Practices (BMP) used to control solid waste and gross pollutants in urban drainage systems which was introduced by Department of Irrigation and Drainage (DID) Malaysia [8]. Main parameters to be considered in designing waste traps are flow rates, maintenance requirements and loading, including the ease of cleaning the waste trap

[9][10]. The main purpose of this paper is to explore users' insights on the factors that contribute to the occurrence of floods caused by clogged drains and to identify user requirements in designing an eco-care smart drainage application. This study needs to be done to determine the culture, social and history of an area because the success of a smartphone application related to solid waste management depends on the local community [11].

Stormwater problems have become more severe due to the increase in urbanisation [12]. Not to mention the increment in the amount of the impervious surface of urban areas produces more stormwater runoff [13].

2. Materials and Methods

The study has been conducted with one FGD with 10 participants. The selected participants are of various age groups, gender, ethnicities, who very observant and concerned about the environment. **Fig. 1** and **Fig. 2** shows the waste trap design concept for open channel respectively used in this study [6].

Fig. 1. Waste trap in the drainage system

The findings were divided into two different sections: who care about drain waste and who does not. Each discussion was audio-recorded and transcribed later. The faces of the participants during the interview were also observed and considered. Then, the data was coded, categorised, and analysed by identifying the importance of the drain waste trap.

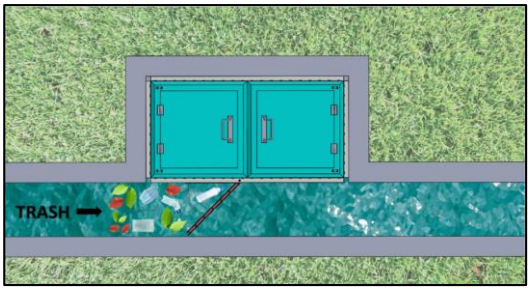


Fig. 2. Top view of the waste trap

3. Results

Characteristics and information of the participants are, as stated in Table 1. The outcome of the Drain Waste Trap Application is divided into six themes – lifestyle and self-monitoring, education and awareness, motivation and commitment, attitude, social support and coaching and technology. These themes emerged from the responses of the survey from the user requirements for the application system.

Table 1. Characteristics of 10 Participants

Characteristics	No.
Age (years):	
21 – 30	5
31 – 40	3
41 – 50	2
Gender:	
Male	6
Female	4
Ethnicity:	
Malay	5
Chinese	3
Indian	2

3.1 Lifestyle and Self-Monitoring

Five participants of this project mentioned that our lifestyle needed to be changed to manage and prevent the problem to occur. In many developing countries, drainage systems are chocked with litter and solid wastes. The people have little knowledge of the effects that can happen monsoon season where the drainage systems and sewer systems are flooded, and water finds its way into the streets and people's homes. Education is therefore essential in providing awareness to the people about the

dangers and causes of floods, and what can be done to minimize its impact on the environment and river. A participant shared why he thought people have littering problems. "Irresponsible people litter in public areas —streets, parks, even by roadsides".

All ten participants agreed for the proper education of preventing littering among the younger generation and also to adults with a sense of humanity and common sense. They shared that they were always on the look-out and were still observing if they spot any littering happening while taking a stroll or even just passing by a road. In the long run, the study shows that commitment in self-management behaviourism is the maximum outcome influencing the success of improved environmental flood problem. Thus managing a proper lifestyle and self-monitoring is essential in taking responsibility for the society in preventing floods caused by littering in clogged drainage systems.

3.2 Education and Awareness

The participants acknowledged the fact that there was less awareness about the waste blockage problem causing drains to be clogged and the factors that cause this problem and with the dangerous water quality caused by pollutants in the open channel flows. Now, there is more awareness about the issue and managing the lifestyle. Presently, there are also more programs such as talks, forums and discussion sessions coordinated by the government as well as the non-government organisations to help educate the public and even expand the use of learning through mobile devices. The participants mentioned that even professionals and higher-ups are thoughtfully educating people regarding the issue of littering, separating wastes according to its' designated bins, what clogged drains can cause to the landslide.

One of the participants mentioned these exact words: "Attending talks expand our knowledge, watching them in a group expands our sense of connection and community. Not all talks may change your life, but you will still find one that can change your life in one way or the other." By attending these talks, participants are more observant and attentive of the way they get rid of their rubbish waste. An obsessive-compulsive disorder (OCD) participant said that there must be more awareness created among the younger generation. The participant mentioned: "Littering is a thoughtless and selfish act. Kids nowadays should be attentive to their social surroundings. If they were, it would not have been a problem for exposure as well as creating awareness that we will be prone to flood if we keep doing what we are doing and don't take any measures to conceal the problem".

3.3 Motivation and Commitment

Several factors provide the participants with being motivated and giving full commitments to the application. They show that motivation and commitment are essential to accept and adapt to this new application. The commitment that we can obtain from this result is that the participants are attentive to what could happen in the future for example, if they do not care about the environment they are prone to have troubles during flash flood especially if their houses are involved. If they know about this application, they would be well prepared if the situation was bound to occur. One of the participants shared that: Everyone in society needs to stay motivated to commit more and care about the environment that we are living now. Without motivation and commitment, our environment is going to get worse, and there is no future here for our children who deserves a clean place to play around and fresh air to breathe in.

3.4 Attitude

Attitude is an essential theme in this research. Without good manners and attitude towards the environment, it may lead to frequent flash flood happening. It is because of clogged drains caused by none other than the society itself. As one of the participants stated before "Selfish and thoughtless acts may trouble us in the future." If no one is wanting to change the attitude about taking care of the environment, the waste itself can spread diseases such as dengue which is a very rampant and worrying disease at the moment. Littering along the road, on the streets or by the bins, those chemicals and wastes can be washed and brought into rivers or sea that pollutes the waterways, and it affects the aquatic life and its' environment.

3.5 Social Support and Coaching

From the result of the survey given out, six out of ten of the participants stated that this project and application needs more exposure and social support to give awareness and coaching to provide a drastic change and outcome towards the clogged drains problem. One of the answers that were received said that it could start from family or even friends. They could begin to the coaching and give timely reminders about the awareness of flood and the uses of the Drain Waste Trap to prevent the society from the normality of littering. One of the participants said: "Only a small number of people know about the exposure or the serious causes of floods and high chemical in water quality. This information could start well from a small group, soon reaching out to a larger group for exposure". Another participant also shared: "This could really save the number of flood damages and even give an outcome of a cleaner environment with better water quality". Other participants shared that it is better to carry out the project first and get a confirmation of the result rather than assuming the possibilities of this application and device. Carrying out awareness about the importance of proper waste disposal from the citizens first was also recorded as an opinion in the survey as it was a better way of using the device efficiently for a faster result. By this way, they learn faster and efficiently to get motivated for a cleaner environment.

3.6 Technology

When technology was mentioned, most of the participants gave favourable responses about the application combined with the trapping device. Although the application will not be connected to the public but instead higher authorities in regarding the update of the trap, involving the use of this modern technology was highly agreed upon as it is a faster responding way to be alert of the device. Instead, the citizens can comply with this project by downloading environmental disaster-related applications from the Google Playstore or the Apple Store to read about the awareness daily. These were a suggestion made by one of the participants of the questionnaire: "I use my smartphone to read the causes of natural disasters and ways to prevent it from getting worse. It is handy since it is on-the-go and you can learn about it anywhere at any time". Since the device application is created to alert the authorities regarding the condition of the traps at any location, the camera devices and detectors need to be applied at every trap station with a link of its own to the application. This is to avoid any confusion of location and detecting the erroneous data in order to do any physical checking when the detection reaches a red zone. A participant shared her viewpoint: "Nowadays, we do not read. Maybe we would if it is on the devices because most of us acquire a smartphone. So information can easily be accessed about this trapping device".

The participants said that the usage of smartphones and Internet or even applications often helps us in many ways. There are many self-care applications about the environment. It is found that some figures self-care applications should be multi-lingual so that many could use it. Also, some mentioned that reminders regarding proper disposing trash could also help this project improve in many ways.

Components of the Proposed Drain Waste Trap Application: The participants highlighted the essential components that they prefer in an application by ranking the system features and components. The priorities of the alerted application components have been divided into two votes: Yes and No. Priorities here means the most important to the least essential components from the participant's viewpoints. **Table 2** shows the percentage of who voted for yes and no for each component. **Fig. 3** illustrates the comparison of each component between the two priorities. Based on the participants' choices and votes, the location and camera instalments stays as the top priority compared to the other components. Then, followed by a chemical sensor as the second most agreed components, damage maintenance, reminder daily check, notifications, awareness exposure for the public, environment information, water quality detector and the daily information with the same results and lastly publicly accessed with the least amount of votes for yes. The same decreasing manner goes on for the votes of no. The lowest result for No, as stated in the chart, is for the components of publicly accessed. Followed by the same votes for daily information and water quality detector, environment information with

Table 2. Participants Priorities on Drain Waste Trap Features & Components.

System Features and Components of Application	Yes (%)	No (%)
Location	100	0
Camera Installments	100	0
Publicly Accessed	20	80
Notifications	60	40
Awareness Exposure	50	50
Reminder Daily Check	70	30
Environment Information	40	60
Water Quality Detector	30	70
Damage Maintenance	80	20
Daily Information	30	70
Chemical Sensor	90	10

40% votes for no, awareness exposure, notification, remainder daily check, damage maintenance, a chemical sensor and lastly with no votes of disagreement, camera instalments and location. It is interesting to note that notification is not being considered as one of the top priority for the components along with the 100% votes for location and camera instalments.

Location, camera instalments and chemical sensors are chosen as the top priorities because it requires to be the essential features in an application. It is personalised for each trap and also to keep track of the condition and situation of the waste trap while monitoring it visually from the application by using the camera instalments. To not get confused about each whereabouts of the trap, they each will be provided with a location system to detect which section will be experiencing damages. Publicly accessed was the least in the priority because the participants find that this information and application is more critical for the authorities compared to the exposure to the public. It is governments duty to keep track of the process. Besides that, daily maintenance and daily reminder check are in the top priorities as the condition of the traps and sensors have to be in good condition in order to run without any errors. Lastly, it is also seen that the daily information and water detector as one of the least priorities before publicly accessed as they are irrelevant information needed.

4. Discussion

These FGD results highlight that majority of the participants would like to use technology as the main component of the application for the device to reduce the occurrence of flash floods due to clogged drainage systems. All the participants agreed with a unanimous vote for the use of location and camera instalments which can be monitored through the application easily. Majority of the society are not aware of the possibilities and how frequent flash floods occur during heavy rain seasons. They tend to bat an eye on the fact that even the slightest action of littering can comply with the phenomenon of natural disasters when they are the ones who can prevent it from happening. It is also known that even when they realise the action is wrong and have been urged about the circumstances, society still would not change their attitude to enhance the beauty and cleanliness of the environment. With the application, as agreed by the participants, this can help increase the value of humanity and form a positive attitude of society. For these reasons, we firmly believe that this application with the assigned features and components will be a stepping stone to prevent flash floods and increment of rubbish surrounding the environment. Chemical sensors will help detect data of pollutants and high chemical reading of the water flow. These can help to increase the water quality in river flows and open water after passing the drainage systems. In order to do this, a sensory detector will be placed at the water level of the channel and will give a reading pop up at the application about the condition of the water. On another note, none of these devices can function well if it is frequently damaged. With this, damage maintenance is essential to always keep the devices in the traps in good condition to allow a smooth process during emergencies. Any failure could show that the application is not well prepared for exposure and will cause more trouble and errors to clear up. With this, damage care is one of the most frequent and agreed components that the applicants chose. Since sometimes damage checks might be neglected, the reminder daily check would be the key to this process.

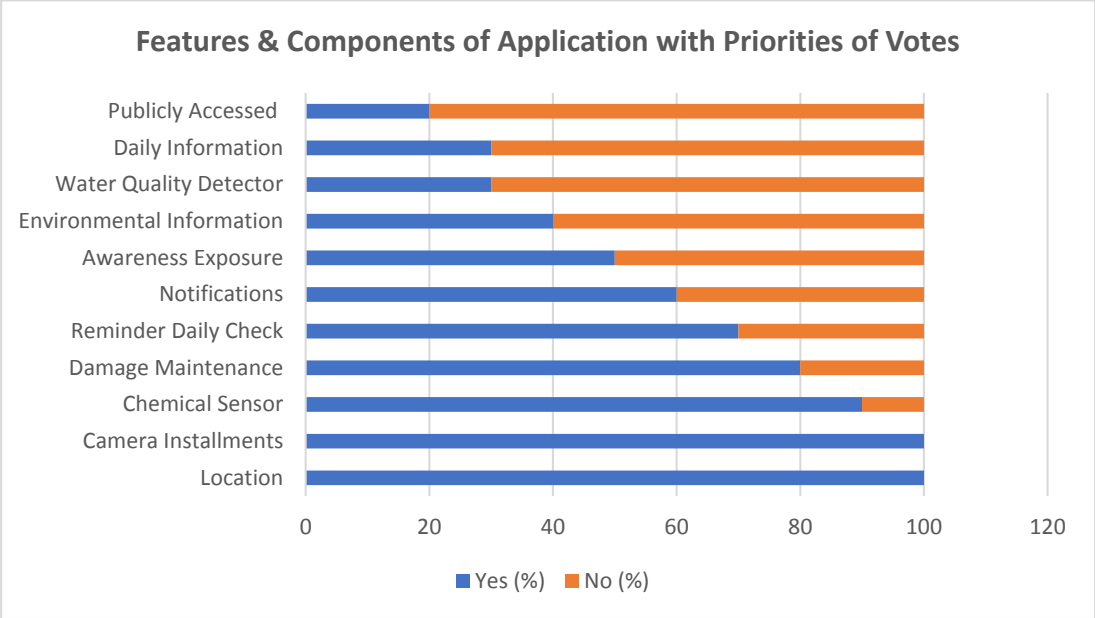


Fig. 3. Comparison between votes (Yes and No) for the system features and components of Drain Waste Trap Application.

Daily, the authorities will be notified to have a physical check to every section location every day in order to prevent any damage from happening. Not only that, in relation to these communications with the authorities, notifications are also to be known as the primary connection. In any circumstances, a signal or notification will alert whenever the weight capacity reaches its limit, when the water quality is too high, or when there is a detection of unwanted living creatures located in the trap. It will also notify whenever a checking is due, or any damage happens. Awareness exposure and environment information could also be categorised together as a way to expand the society's knowledge of environmental issues as well as natural disasters. This information and features will help them to keep track of their surroundings and manage their littering more efficiently. As for the idea of making this application publicly accessible, they did not show much interest as they said that the idea is only suitable for the use of the authorities regarding the Drain Waste Trap. If a person is not interested, then it would be difficult to change their opinion. Despite getting positive feedback on the idea of developing the application, about 20% of the participants showed less interest in the existence of the application. These are either because they are not technology savvy and do not own a smartphone or because they do not care about the surrounding issues that have been going on.

5. Conclusions

Collecting the feedback on the idea of developing a trash trap application and their requirements for the application is essential to understand the needs and expectations of the target users so that a useful trash trap application can be designed for them to keep track of the drainage system around them independently. Even though our Focus Group Discussion (FGD) study is based upon small sample size, it provides us valuable insight into the participants' lifestyle, awareness on flood by littering, and how they get motivated to change their behaviour. These insights help much to determine their expectations towards using technologies such as IoT as opportutnies to managing waste in taking care of the environment [14]. The Focus Group Discussion study results support our hypothesis that a trash trap application is a suitable tool in assisting the trash trap in preventing flood from happening. Our study also confirms that usage of technology is low among older people and to reduce the

digital devide among this individuals are highly recommended [15].

References

[1] R. A. Allison, T. A. McMahon, and F. H. Chiew, *Stormwater gross pollutants*. CRC for Catchment Hydrology, 1997.

[2] G. E. Louis, 'A Historical Context of Municipal Solid Waste Management in the United States', *Waste Manag. Res.*, vol. 22, no. 4, pp. 306–322, Aug. 2004, doi: 10.1177/0734242X04045425.

[3] K. Palmer, H. Sigman, and M. Walls, 'The Cost of Reducing Municipal Solid Waste', *J. Environ. Econ. Manag.*, vol. 33, no. 2, pp. 128–150, Jun. 1997, doi: 10.1006/jeem.1997.0986.

[4] J. Madhani, S. Madhani, and R. Brown, 'A literature review on research methodologies of gross pollutant traps', in *Proceedings of the First International Conference on Engineering, Designing and Developing the Built Environment for Sustainable Wellbeing*, C. J. L. Cowled, Ed. Australia: Queensland University of Technology, 2011, pp. 145–150.

[5] R. A. Allison, T. A. Walker, F. H. S. Chiew, I. C. O'neill, and T. Mcmahon, 'From roads to rivers: gross pollutant removal from urban waterways', 1998, Accessed: Oct. 24, 2020. [Online]. Available: /paper/From-roads-to-rivers%3A-gross-pollutant-removal-from-Allison-Walker/bb1c577f15f661c2dda32a58f25e3c5b7fb43988.

[6] S. Hassan and N. Abd Rahman, 'Study of Effectiveness for Rubbish Traps in Open Channel Flow', Aug. 2007.

[7] M. K. F. B. Sarimin *et al.*, 'The Effectiveness of Proposed Design Rubbish Trap System'.

[8] Government of Malaysia Department of Irrigation and Drainage, *Urban stormwater management manual for Malaysia*, 2nd ed. 2012.

[9] N. A. Naamandadin, N. Nordin, W. Mustafa, and N. Z. Zainol, 'Stress Analysis of Urban Solid Waste Trap by Using Finite Element Analysis (FEA)', 2019.

[10] P. L. Brezonik and T. H. Stadelmann, 'Analysis and predictive models of stormwater runoff volumes, loads, and pollutant concentrations from watersheds in the Twin Cities metropolitan area, Minnesota, USA', *Water Res.*, vol. 36, no. 7, pp. 1743–1757, 2002.

[11] A. Mavropoulos, M. Tsakona, and A. Anthouli, 'Urban waste management and the mobile challenge', *Waste Manag. Res.*, vol. 33, no. 4, pp. 381–387, Apr. 2015, doi: 10.1177/0734242X15573819.

[12] J. D. Miller and M. Hutchins, 'The impacts of urbanisation and climate change on urban flooding and urban water quality: A review of the evidence concerning the United Kingdom', *J. Hydrol. Reg. Stud.*, vol. 12, pp. 345–362, Aug. 2017, doi: 10.1016/j.ejrh.2017.06.006.

[13] A. Müller, H. Österlund, J. Marsalek, and M. Viklander, 'The pollution conveyed by urban runoff: A review of sources', *Sci. Total Environ.*, vol. 709, p. 136125, Mar. 2020, doi: 10.1016/j.scitotenv.2019.136125.

[14] T. Anagnostopoulos *et al.*, 'Challenges and Opportunities of Waste Management in IoT-Enabled Smart Cities: A Survey', *IEEE Trans. Sustain. Comput.*, vol. 2, no. 3, pp. 275–289, 2017, doi: 10.1109/TSUSC.2017.2691049.

[15] N. G. Choi and D. M. DiNitto, 'The Digital Divide Among Low-Income Homebound Older Adults: Internet Use Patterns, eHealth Literacy, and Attitudes Toward Computer/Internet Use', *J. Med. Internet Res.*, vol. 15, no. 5, p. e93, 2013, doi: 10.2196/jmir.2645.

Exploration of Web Links and Network Collaboration Between University's Libraries in Southeast Asia

Su Elya Namira Mohamed, Nor Azman Ismail
School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia
Skudai, Malaysia
[e-mail: senamira2@live.utm.my, azman@utm.my]
*Corresponding author: Su Elya Namira

Abstract

Webometric analysis and social network analysis is a new way in indicating the performance of a website and their behaviour in the web space. The institutional library is said to be one of the factors a university to be visible in the web space. This study aims to map and explore the relationships initiated on the web between the institutional library of the top 10 universities in Southeast Asia. Top 10 university in Southeast Asia was chosen for sample selection based on the QS-World University Ranking. Data on each institutional library's website were fetched using webometric techniques (backlink), and a thorough calculation was generated for social network analysis. The result shows that the backlink only frequently happens in the same country which can be considered as a geopolitical relation. The result also shows that Universitas Indonesia Library has the highest value of site size which contributes to the biggest influence in the search engine. Nevertheless, Mahidol University Library and Knowledge Center takes the 'mediation' role in the network but lacks public acknowledgement on the web by hyperlinks. In sum, exploring the network of the libraries of the top 10 universities in Southeast Asia shows that the institutional library has a lack of collaboration in the web space in terms of providing and using resources. Despite their university ranking status, their institutional library has lack of acknowledgement in the network. Reinforcing web integration policies and increasing networking through the exchange of hyperlinks is advisable. The internet could thus represent international cooperation in academia and help legitimize and enhance the exposure of the existing networks of collaboration between universities in Southeast Asia.

***Keywords*:** Webometric, Social Network Analysis, Library, Southeast Asia University

1. Introduction

Libraries have shifted their role to being providers of access-based digital information resources from being keeper of the collection-based traditional information resources which they have to publicize and ease the access to the variety of information through their websites to their users'. The studies of the existence of the library website of the university on the worldwide web typically measure the quality of their web domain in the same way as measuring the universities' educational and research results.

Since the academic web domain was a means of improving and promoting a university's research, science and educational skills, we are taking an approach to investigating the relationship between its institutional library on the web. To be able to visible in the world wide

web, the web domain needs to assure their acceptance in the network by producing more resources that can be used by other user. One of the ways to be visible in the web space is through the institutional libraries which its contribution in producing educational resources is significant.

According to Aminpour et .al, the ranking of universities worldwide has long been calculated by their webometric features and has had a significant impact on the web domain of the university [1]. In the World Wide Web platform, webometric is the most reliable technique in determining the ranking of a website. It focuses primarily on web-based communication units like pages, web domains and hyperlinks as a means of understanding new scientific activities [2]. Also, according to Ortega et .al, one of the techniques used as a web indicator is the Social Network Analysis technique which is it helps to study the main characteristics of webspace and measure the structural relationship of a website with its surrounding [3].

Webometrics studies focuses the studies mainly on the performance of the academic web domain, and this is because the academic institution such as college and universities are the institution that are stable and well-defined on the web [4]. The institutions are selected on the basis of the QS-World University Ranking in the Southeast Asia region which focuses only on the Top 10 university (**Table 1**).

The aim of this study is to explore the behaviour of the institutional library's website. We believe that mapping the relationships between the institutional library that these institutions have formed on the internet will reveal the key players in this network, the least connected institutions and the impact of each institutional library in the Southeast Asian academic institution which also can be a guide for the institution to plan their web content strategy in the future.

Table 1. University and its corresponding library

University	Library
National University of Singapore (NUS)	NUS Libraries
Nanyang Technological University (NTU)	NTU Library
University of Malaya (UM)	University of Malaya Library
Universiti Putra Malaysia (UPM)	Perpustakaan Sultan Abdul Samad
Universiti Kebangsaan Malaysia (UKM)	Perpustakaan UKM
Universiti Sains Malaysia (USM)	Perpustakaan Hamzah Sendut
Chulalongkorn University (CU)	Chulalongkorn University Library Information Network
Universiti Teknologi Malaysia (UTM)	UTM Library
Mahidol University, Thailand (MI)	Mahidol University Library and Knowledge Center
Universitas Indonesia (UI)	Universitas Indonesia Library

2. Methods

In general, the study was focused on webometric and social networking analysis techniques. Webometric concept is that the analysis of the quantitative aspects of the construction and the use of data tools, structures and technologies on the web drawing is focused on bibliometric and informetric approaches [5]. In a webometric analysis, backlink contributes 50% of the calculation and site size contributes 20% of the calculation. In this paper, referring domain data was collected to map the number of hyperlinks shared between two websites. Referring domain is referred to a domain that people came from when visiting a website. For this project, data extraction for the referring domain is done by using the Ahref backlink crawler.

The methodology of social network analysis (SNA) was applied to the data in order to better understand and visualize the connections between institutions. Social network analysis work as an indicator to measure the main characteristic of a network. It allows us to compare and to get an insight into how a network is structured. Three of the technique used in this research is:

Degree: Is the number of node-adjacent edges. It is used to display how deeply an entity is related to its counterparts. The inlink/indegree and outlink/outdegree is calculated to make the evaluation.

Betweenness: It calculates the degree of intermediation based on the number of shortest paths that pass through a node. To select the shortest path and calculate the betweenness centrality, Dijkstra's algorithm is used. This measure makes it possible to detect hub or gateways linking different web sub-networks [6].

Closeness: This calculates the average distance with each node in the network in a number of

clicks of a node. It is focused on a website's proximity to the rest that can be a good indicator for studying processes of infection and flows of data. It also indicates a website's level of visibility in the navigation process.

The way to represent the graph is by applying the Spring algorithm in force-directed layout which is the most flexible algorithm in calculating the layout of a simple undirected graph. To create the visualization, Cytoscape tools were used as it is the most convenient tool for interpreting network analysis.

3. Results

Network analysis using Cytoscape tools presents 10 institutional libraries grouped into a single component. **Fig. 1** shows the link relationship between each institutional library. The value of the backlink is interpreted in color code and the size of the node is representing the size of the site in the webspace.

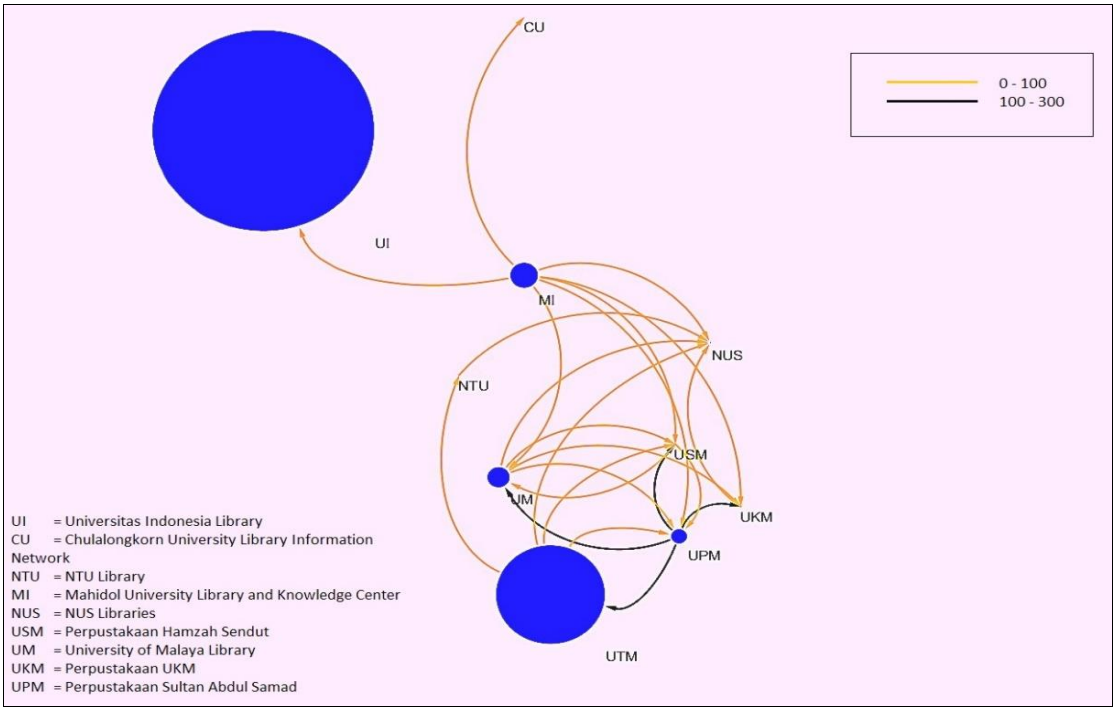


Fig 1. A Web Environment Network for Top 10 University in Southeast Asia

Table 2. Centrality measures

Institutions	Indegree	Institution	Outdegree
Perpustakaan Hamzah Sendut	349	Perpustakaan Sultan Abdul Samad	1059
University of Malaya Library	295	University of Malaya Library	74
UTM Library	289	Mahidol University Library and Knowledge Center	20
Perpustakaan UKM	186	UTM Library	11
NUS Libraries	34	Perpustakaan UKM	9
Perpustakaan Sultan Abdul Samad	33	Perpustakaan Hamzah Sendut	9
NTU Library	2	NTU Library	6
Universitas Indonesia Library	2	NUS Libraries	0
Chulalongkorn University Library Information Network	1	Chulalongkorn University Library Information Network	0
Mahidol University Library and Knowledge Center	0	Universitas Indonesia Library	0

Perpustakaan Sultan Abdul Samad has the largest value of outlink compare to other libraries which shows that this library is frequently referring to other libraries for their resources. Even though Universitas Indonesia Library has the highest value of site size which shows that the web domain has the biggest influence in the search engine, it does not reflect their contribution in providing good resources that can be referred by other institutions. Also, the institution is not showing a good contribution in developing the engagement with other institutions among the top 10 universities as we can see that it only has a single connection within the region. The connection between each institution frequently occurs within a country which shows that the geopolitical relation gives some impact on developing institutional engagement.

3.1 Degree Centrality (Indegree and Outdegree)

An evaluation of the statistics found that the organization with the highest number of links is not the largest provider. In fact, the links come from the same institution with a large number. Both the Indegree (a link receives from other domain) and Outdegree (a link given to other domain) values for the top 10 institutions are shown in **Table 2**.

Perpustakaan Hamzah Sendut, ranked in the first place in receiving the inlinks but, in giving the outlink, this institution ranks in the 6th place. The University of Malaya Library maintains its second place in both indegree and outdegree. This shows that the organization has a well-balanced interconnection by identifying and acknowledging its partner. On the other hand, Chulalongkorn University Library Information Network becomes the least active institution in the network as it has the least amount of indegree and outdegree. While their collaborators are widely recognized, Perpustakaan Sultan Abdul Samad and Mahidol University Library and Knowledge Center are not equally recognized as creating a void in how this partnership is represented on the internet.

3.2 Betweenness Centrality

According to the betweenness centrality calculation done (Table 3), Mahidol University Library and Knowledge Center seemed to be the first with a higher capacity to draw partners and serve as a gateway for linking various institutions to the center of the network even though it is not one of the most highly connected institutions on the network. It shows that this institution is important in terms of communication, transportation or transaction and connection between other libraries which it takes the 'mediation' role in the network.

This network includes the institution from several countries which is Malaysia, Singapore, Thailand, and Indonesia. Those collaboration involving these countries is still not reflected by hyperlinks on the web as the network is majority conquered by a Malaysian institution. In Malaysia regions, an institution is still recognizing more Malaysian partners as key players as they provided the links more to Malaysian institutions compared to the institution from the other countries.

Interestingly, the institution from Thailand stands out as link providers to the other institution from other countries and finally taking the mediation role in the network.

3.3 Closeness Centrality

Closeness centrality indicates a more central position for a certain node. In this case, Perpustakaan Sultan Abdul Samad has the lowest closeness value which shows that this institution is directly connected or just a hop away from its collaborators (**Table 4**). The relationship develops among the institution is most likely within the same country. The enhancement of the engagement between the institution within a country is said to be more efficient and effective to make the country to be well-known. This also can be related to geopolitical relations. NTU Library has the highest value of closeness centrality which indicates that this institution has a number of hops or connections it needs to take to connect to distant others in the network. Closeness centrality also can be remark as a measure of how long it will take to spread information from a node to another which stands for the convenience and ease of connection between the nodes. It easier to spread information to the institution within the same country compare to the institution that out of the country. From ten institution, there is only seven institutions that able to come out with a closeness

centrality value. The other three institution does not have enough connection in the network to be included in the calculation of closeness centrality.

Table 3. The institutional library by their betweenness centrality

Country	Institution	nBetweenness
Thailand	Mahidol University Library and Knowledge Center	18
Malaysia	Perpustakaan Sultan Abdul Samad	9.5
Malaysia	University of Malaya Library	9.3
Malaysia	UTM Library	7
Singapore	NUS Libraries	5.5
Malaysia	Perpustakaan Hamzah Sendut	5
Malaysia	Perpustakaan UKM	3.1
Singapore	NTU Library	0
Thailand	Chulalongkorn University Library Information Network	0
Indonesia	Universitas Indonesia Library	0

Table 4. The institutional library by their closeness centrality

Country	Institution	nCloseness
Malaysia	Perpustakaan Sultan Abdul Samad	0.0085
Malaysia	University of Malaya Library	0.12
Thailand	Mahidol University Library and Knowledge Center	0.45
Malaysia	UTM Library	0.82
Malaysia	Perpustakaan UKM	1
Malaysia	Perpustakaan Hamzah Sendut	1
Singapore	NTU Library	1.5

4. Discussion

The goal of the present study was to map and explore the relationships initiated on the web between the institutional library of the top 10 universities in Southeast Asia by applying the webometric analysis and social network analysis technique. The investigation apparently shows that there is not much collaboration between the institution. The relationship also was affected by the geopolitical relation which they prioritize the institution within the same country. The results showed that Malaysian institutional libraries such as Perpustakaan Sultan Abdul Samad, University of Malaya Library and UTM Library are the most connected on the web and have a higher capacity to attract hyperlinks. On the other hand, NUS Library which is from a well-known institution have a lack of recognition on the web by means of hyperlinks. Many researchers have done the web analysis to investigate the visibility of the chosen website in the world wide web network applying the webometric analysis technique. As webometric analysis is involved with the link analysis, it can be manipulated, and it can be inaccurate and to make the results reliable, the social network analysis technique was used. According to Ortega et .al, one of the techniques uses as a web indicator is the Social Network Analysis technique which is it helps to study the main characteristics of webspace and measure the structural relationship of a website with its surroundings [3].

5. Conclusion

In conclusion, the university library's presence in the web space does not affected by the university ranking in the world. In fact, the institutional library investigated has lack of collaboration in the web space in terms of providing and using resources. Therefore, reinforcing web integration policies and increasing networking through exchange of hyperlinks is advisable. The internet could thus represent international cooperation in academia and help legitimize and enhance the exposure of the existing networks of collaboration between universities in Southeast Asia.

References

[1] Aminpour, F., Kabiri, P., Otroj, Z., & Keshtkar, A. A. (2009). Webometric analysis of Iranian universities of medical sciences. Scientometrics, 80(1), 253-264.

[2] Ortega, J., Aguillo, I., Cothey, V., & Scharnhorst, A. (2007). Maps of the academic web in the European Higher Education Area—an exploration of visual web indicators. Scientometrics, 74(2), 295-308.

[3] Ortega, J. L., & Aguillo, I. F. (2011). Social Network Tools for the Assessment of the University Web Performance. In Social Media Tools and Platforms in Learning Environments (pp. 185-201). Springer, Berlin, Heidelberg.

[4] Maharana, R. K., Panda, K. C., & Sahoo, J. (2012). Web impact factor (WIF) and link analysis of Indian Institute of Technologies (IITs): A webometric study.

[5] Björneborn, L. (2004). Small-world link structures across an academic web space: a library and information science approach. Royal School of Library and Information Science. Department of Information Studies.

[6] Aguillo, I. (2010, March). Web, webometrics and the ranking of universities. In Proceedings of the 3rd European Network of Indicators Designers Conference on STI Indicators for Policymaking and strategic decision, CNAM, Paris (to appear, 2010).

[7] Chakravarty, R., & Wasan, S. (2015). Webometric analysis of library websites of higher educational institutes (HEIs) of India: A study through google search Engine. DESIDOC Journal of Library & Information Technology, 35(5).

[8] Dictionary, m. (2018). metrics Meaning in the Cambridge English Dictionary. [online] Dictionary.cambridge.org. Available at: https://dictionary.cambridge.org/dictionary/english/metrics [Accessed 2 Apr. 2018].

[9] Dictionary, w. (2018). web Meaning in the Cambridge English Dictionary. [online] Dictionary.cambridge.org. Available at: https://dictionary.cambridge.org/dictionary/english/web [Accessed 2 Apr. 2018].

[10] Harries, G., Wilkinson, D., Price, L., Fairclough, R., & Thelwall, M. (2004). Hyperlinks as a data source for science mapping. Journal of Information Science, 30(5), 436-447.

[11] Holmberg, K. (2009). Webometric network analysis: Mapping cooperation and geopolitical connections between local government administration on the Web.

[12] Jasani, B. M., & Kumbharana, C. K. (2014). Analyzing Different Web Crawling Methods. International Journal of Computer Applications (0975-8887), 107(5).

[13] Kenekayoro, P. (2014). Collaboration between UK Universities: A machine-learning based webometric analysis (Doctoral dissertation).

[14] Li, X. (2003). A review of the developmenz and application of the Web impact factor. Online Information Review, 27(6), 407-417.

[15] Minguillo, D. (2014). Mapping R&D support infrastructures: A scientometric and webometric study of UK science parks. arXiv preprint arXiv:1407.7250

[16] Ng, A. Y., Zheng, A. X., & Jordan, M. I. (2001, September). Stable algorithms for link analysis. In Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval (pp. 258-266). ACM.

[17] Ortega, J. L., & Aguillo, I. F. (2008). Visualization of the Nordic academic web: Link analysis using social network tools. Information Processing & Management, 44(4), 1624-1633.

[18] Ortega, J. L., & Aguillo, I. F. (2009). Mapping world-class universities on the web. Information Processing & Management, 45(2), 272-279.

[19] Rohrer, R. M., & Swing, E. (1997). Web-based information visualization. IEEE Computer Graphics and Applications, 17(4), 52-59.

[20] Thelwall, M. (2001). Extracting macroscopic information from web links. Journal of the Association for Information Science and Technology, 52(13), 1157-1168.

[21] Thelwall, M. (2012). A history of webometrics. Bulletin of the Association for Information Science and Technology, 38(6), 18-23.

[22] Thelwall, M., & Harries, G. (2004). Do the web sites of higher rated scholars have significantly more online impact?. Journal of the Association for Information Science and Technology, 55(2), 149-159.

[23] Thelwall, M., Vaughan, L., & Björneborn, L. (2005). Webometrics. ARIST, 39(1), 81-13.

[24] Yang, B., Liu, Z., & Meloche, J. A. (2010). Visualization of the Chinese academic web based on social network analysis. Journal of Information Science, 36(2), 131-14

Replay attack to DRM in music streaming service on Android

Sunghee Lim, Jaehyeok Han and Sangjin Lee
Department of Information Security, Korea University
Seoul, South Korea
[e-mail: shee2101@korea.ac.kr, one01h@korea.ac.kr, sangjin@korea.ac.kr]
*Corresponding author: Sangjin Lee

Abstract

Music streaming companies are applying DRM technology for preventing the duplication and leakage of their contents. Despite the efforts, leaks are still being made through various services and distribution channels. For this reason, it is important to find and block these spill routes. We discovered that music streaming files could be leaked by using structural problems that occur in MediaPlayer framework, which includes an API that supports playback of DRM-protected material. In this paper, we proposed methods to bypass DRM in music streaming services. Also, a case study was conducted on well-known Korean streaming applications. We then present methods to prevent leaks of the contents in the perspective of digital forensic investigation.

***Keywords*:** DRM, Bypassing, Stream-ripping, MediaPlayer, Android, Proxy Server, Hooking

1. Introduction

Many users in pirate sites made a lots of visits to illegally access music and to download or stream movies. Stream-ripping is the act of undermining the copyright of the copyright holder. Once distributed through stream-ripping, it is easy to copy due to the nature of digital media and can easily lead to a secondary leakage accident because mass distribution is possible through a web hard site at once. Therefore, a business operator providing digital contents such as books and music records has applied DRM, which includes an encryption process, so that it is possible to safely manage the contents.

In the case of music content services, users can download and then use them, but they mainly provide content through streaming. Service delays by action to protect content must be minimized, since the streaming method must provide services in real time. Contents can be protected by providing content in the service provider's unique file format, such as DFC (DRM Contents File), but it is used limitedly, and most services manage DRM content by configuring a proxy server inside the device.

In this paper, we analyze the application flow targeting 6 music streaming applications, and present an attack model that can bypass DRM with a replay attack using a proxy server. Then, we showed that the replay attack was successful on Android phones. Also, we present counter-measures to be supplemented in streaming service based on the scenario using replay attack.

2. Background and Related Works

The working process for protecting and managing digital content with DRM is called packaging. Packaging uses encryption to prevent illegal copying, and methods such as watermark or fingerprinting that insert owner information so

This work was supported by Institute of Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No.2018-0-01000, Development of Digital Forensic Integration Platform)

that the original source can be traced even if copied to the outside.

In order to verify whether content protection is working well in DRM, various attack techniques are being studied along with the development of related protection techniques. [1] experimented to obtain a video chunk decryption key transmitted over HTTPS with MITM attack in order to make DRM-encrypted video playable. [2] presented an attack model using android WebView component in E-book application. [3] identified music content and metadata exposure in http and https connections through packet sniffing and MITM attacks, and applied SSL pinning to prevent MITM in https connection.

API supporting DRM decryption is required to view DRM-protected content, but it is not provided in the Android operating system. Therefore, the android application creates a proxy server, then releases DRM from the proxy server and serves the music locally as shown in **Fig. 1**. In the process of serving music, MediaPlayer in the Android framework, is provided as a class to play raw data existing on the file system and media contents files provided in the form of data streams. MediaPlayer is to play voice or video files. The data stream received through the internet can be played DRM-encrypted file with the MediaPlayer API.

3. Replay Attack to bypass DRM

3.1 Operation in music streaming services

After a user requests a music contents, the application checks whether the login is successful and the subscription is in progress. Depending on the results of the authentication, streaming services are available via proxy server. The process to play music file is as follows. First, the application configures a proxy server and requests user authentication to the contents vendor server (Step 1). According to the authenticcation result, the path to the DRM-encrypted file is obtained (Step 2). Then, transfer the obtained path to the proxy server (Step 3). The proxy server requests a music file to the content vendor server (Step 4), and receives a response to the DRM-encrypted content (Step 5). The received file(streaming format) is played in the MediaPlayer (Step6).

The application downloads DRM-encrypted files, which can be played in whole or only one

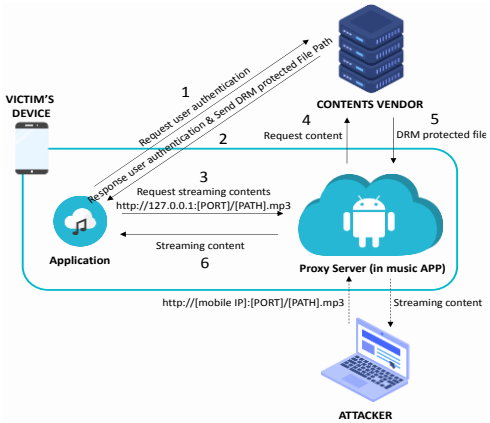


Fig. 1. Replay attack model using proxy server in the music streaming service

minute according to the license. In order to decrypt DRM-encrypted files, some applications use the proxy architecture style and open their own proxy server to decrypt. This decrypted file is played in MediaPlayer through http protocol. In this process, a connection between the client and the proxy server is created.

We try to use the structural vulnerability of providing streaming content through proxy servers to obtain content without DRM decryption. In this case, the content file can be taken away without any attempt to break the DRM. Even if there are different DRM method, this attack can be applied to any application that uses MediaPlayer. This type of communication has structural weaknesses and is vulnerable to port scanning and automated brute force attacks. We have identified the process of applying any port in the form "http://127.0.0.1:%d/" and then adding variable value to URL parameter. To connect to the proxy server, we found strings to be parameter and collected addresses for access to the streaming contents provided by the local proxy.

Using this, we checked the process of handling the URL in real time. While the session is maintained, it uses a fixed port number and has an identifier for each music contents. When a DRM decrypted music file is played in streaming form, we hooked the function and logged. Simply access the URL obtained in the process, and we can collect decrypted music files. This can also be seen by capturing packets that access the proxy server from a routed device. This method can be used even if a debugging framework, such as Frida, is not available and can be referred to by an attacker to find a streaming path.

The file received in Streaming form is mpeg format and that is not applied DRM. In such cases, after the copyrighted file is circulated, tracking copyright holders and buyers, illegal distributors of content become difficult. Also, once the file is exposed in streaming form, the user can use the file without any time limit.

3.2 Experiments of exploit for PoC

Attacker can be accessed through the proxy server of the target user's device since the application operates a proxy server. The attacker can obtain the URL requested in step 3 by packet sniffing (or malicious disseminators). then attacker can obtain the same streaming file by replaying http requests. Then it is available to identify the streaming path. This replay attack is possible because the content vendor server does not authenticate and does not block a connected session once. If the application uses proxy server, the music file can be distributed to many users who are not licensed. These attacks tend to have a low risk of attack due to frequent IP changes by cellular and Wi-Fi. but once exposed, it is critical. The experiment was conducted after installing Frida-server with rooted SHV-E300S installed Android version 7.1.2. We analyzed 6 applications which are provides music streaming contents and released from August 2020.

Comparing the results of the attack before and after logging in, all downloaded music files were playable, with only the playable time was different. Several applications using a proxy server are limited to use of a proxy server when providing streaming content. Another application used a proxy server only if the device played DRM encrypted files that had already been downloaded without the Internet. We measured session timeout for up to 60 minutes and confirmed that session was maintained for more than an hour or less than 30 minutes. Some uses 80 port directly for connection with a server.

4. Countermeasures for Replay Attack

In this section, we introduce the four methods to prevent the attack presented above and the method to track the distributor when streaming contents are exposed by attack.

***S_1*. Use token.** By using a one-time token when requesting to the proxy server, you can prevent the attacker from reusing the request. Alternatively, if a sequence number is assigned to the request packet and the count is increased each time a request comes to the server, content provision can be restricted based on the count value.

***S_2*. Use timestamp or session timeout.** It is a method that puts timestamp in content request packets. When the application accesses the proxy server, the server checks timestamp and blocks it. The threshold of the timestamp is set as the maximum time value in case of replaying the same song (e.g. replay, failed to play, etc.).

***S_3*. Port randomization.** This is a method of changing the port number after the first streaming contents request to the proxy server. The port number must be changed to a random value.

***S_4*. Restriction of access to proxy server through authentication.** There is not enough management on disconnecting session after requesting the content. There are two ways to disconnect session. One method is to restrict when the same request comes to the proxy server, and the other is to block when a request comes from an external device. For this, device's unique ID or vendor server need to be authenticated at the first time, then proxy server can block all TCP connections for unknown source.

5. Conclusion

We have contributed to a stronger mobile security environment by providing mobile application developers with a systematic approach to finding design weaknesses and vulnerabilities and formulating fixes. We demonstrated how to use the attack model so that application developers can discover vulnerabilities in streaming applications and understand how to exploit them. We also provided four recommendations that developers can use to enhance the security of music streaming contents.

References

[1] D'Orazio, Christian, and Kim-Kwang Raymond Choo. "An adversary model to evaluate DRM protection of video contents on iOS devices." Computers & Security 56, 94-110, 2016.

[2] Lee, Sangmin, et al. "Bypassing DRM protection in e-book applications on Android." 2018 International Conference on Information Networking (ICOIN). IEEE, p.194-198, 2018.

[3] Jung, Woo-sik, Hyun-gyu Nam, and Young-seok Lee. "Analysis and Prevention of Contents Exposure in Music Streaming." KNOM Review 21.2, 10-17, 2018

CONTEXT-AWARE SYSTEM TO LOCATE OBJECTS USING GOOGLE AWARNESS API

Maheen Gul Mastoi, Dr. Isma Farah Siddiqui*, Shafiya Qadeer Memon, Mariam Jawaid Memon
Department of Software Engineering, Mehran University of Engineering and Technology, Jamshoro, Pakistan
[e-mail: maheen_mastoi@hotmail.com, isma.farah@faculty.muet.edu.pk]
*Corresponding Author

Abstract

In today's hectic world that we live in, we all have that habit of misplacing things time to time. Normally people spend hours finding their required stuff which not only makes it frustrating, but also time consuming. Thus having an application that can easily lead us to things we're trying to find will not only be productive for us but a savior in many situations. The FoundIt is one such context-aware android application that would help users to easily organize and find their objects placed in cupboards or cabinets of their homes/offices without any trouble. The application makes use of beacons which are added to the cupboards or cabinets, these beacons let the user(s) know about the things placed inside the closet. The FoundIt also offers a search feature using which the users can easily find out if they've the required item available in the closets and would also specify the exact place of the item. This application is useful in places like homes, offices, schools e.t.c for people (like nannies, caretakers,members of the house, office employees e.t.c) who aren't much familiar about the placement of the required things or often forget where the necessary things are located.

Keywords Google Awareness API, Bluetooth Beacon, Contect-aware Systems

1. Introduction

The integration of mobile computing and low-energy bluetooth devices has helped create such smart applocations to ease our day-to-day tasks. The Bluetooth Low Energy or BLE devices help to receive and broadcast signals in short distance range[1]. The beacon thus pushes notification of the items inside the cupboard, helping the user to find required items using their smart phones. The user needs to have phone's bluetooth enabled and item-locator application downloaded to find his items.

2. Bluetooth Beacon

The Low Energy devices, Beacons, are the part of Blutooth protocol V4. Bluetooth technology generally operates at a short wavelength ranging from 2.4GHz to 2.485GHz. BLE comprises of two parts: a receiver (smartphone with required app) and broadcaster (small beacon device). The small beacon devices transmit the same packet of information in every n milliseconds. The beacons use signals to broadcast UUID (Universally Unique Identifier) and the Android device uses that UUID to trigger and display the required message.[2][3]

3. Bluetooth Applications

Many applications have been taking in beacon technology to enhance the functionalities over the last few years. Be it Apple, Macy's, Google or Virgion Atlantic, these major brands and companies have been using BLE technology to enhance customer experience and marketing strategies.[4] From online shopping, discount

offers or promotions notifications, checking-in at games to in-flight activities and currency exhange offers; beacons are now used to find ways to enhance retail environment experience for consumers. Other than; for marketing purpose, beacons are also used in many applications to automate tasks on user devices, take attendance in classroom or to remind employees about their day-to-day tasks

4. Proposed System Architecture

The **Fig. 1** shows a beacon attached to a cupboard. The beacon transmits continuous signals which are detected by nearby android devive. The device with smart bluetooth enabled and application installed in it gets notified about the items placed inside the cupboard, this make it easier for the smartphone user to track his items. Bluetooth beacons, as shown above, are small wireless hardware transmitters that broadcast signals to nearby devices. Beacons can be placed any where such as on an office desk, in a classroom or on a closet to make use of the in-door navigation and other facilities provided by this small BLE device. The proposed system comprises of the things mentioned below:

a) The proposed system includes an Android based mobile application with bluetooth signal scanning functionalities.

b) Beacons are placed at various places such as for office cabinets, home cupboards/cabinets e.t.c which will broadcast the items placed inside it.

c) This system uses a Context-Aware API that helps to detect the Beacon signals easily; making an efficient use of it.

d) It also helps user to search the required item by using the search functionality in the Application.

5. Implementation

A. Android Devices

The Android devices with a Bluetooth LE chipset and Android platform (version 4.3 or above) can detect beacons and thus can work using this system.[5] The android devices must also contain Google Play Services in order to work with the Awareness API.[6]

B. Context-Aware API

The target Android device should be able to work with Context-Aware API, If the target device doesn't have play services, the proposed system won't work on it. Sonce the Google Awareness API itself is a part of Google Play services.[7]

C. Beacons

The small physical devices, beacons, are to be placed on places where needed. These small devices will enable the bluetooth devices around to receive or broadcast small static piece of data while constatntly transmitting radio signals. [8] To send signals as a beacon, Android version 5 (or above) and firmware supporting BLE peripheral mode are needed.

D. Implementation Requirements

The proposed system is not restricted to single-function devices and it can be incorporated in any Android device with Bluetooth LE feature. The Android device needs to have Google Play Services incorporated in it and version of 4.3+ to work with beacons. The beacons are to be attached to the places from where the information is required to be transmitted to the devices. The Android device needs to have its Bluetooth enabled and required application installed to work with this system.[9][10]

6. Conclusion

This paper provides an overview of how a system which uses Bluetooth LE device i-e beacon and Google's context aware API to effectively search for the required item(s) using smart phone. The system is based on Android devices which must have Google Play Services in order to work with Google's Awareness API.

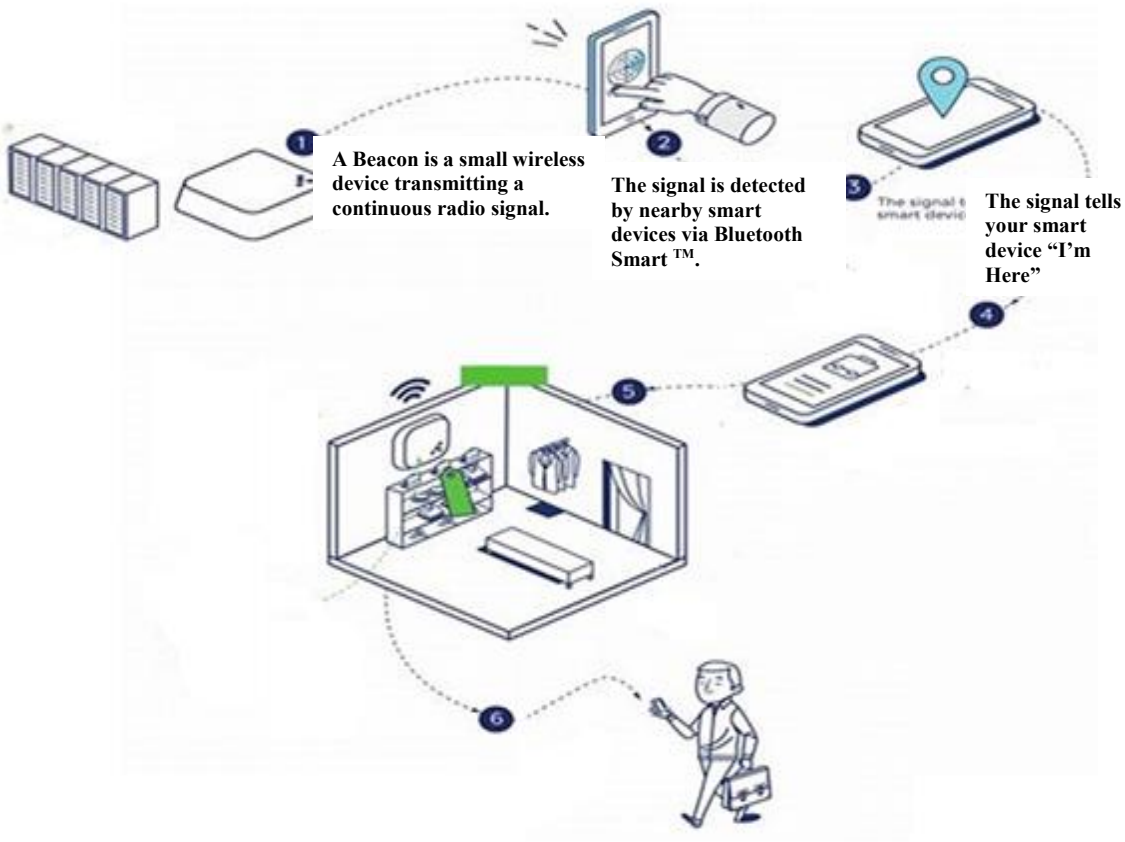


Fig. 1. Interaction between Beacon and SmartPhone

The Awareness API has the capability of sensing user's location and context information in order to create powerful context-based applications that may be used in several marketing, busuiness and many other such applicatons. For the system to be implemented, the smart phone must have bluetooth enabled and the required application installed in order to serve the required purpose. For future work, the system can be motified to implement automatic installation of items' information into the beacon rather than storing the information manually. If any items are removed or added inside the cupboard/cabinet, the application would be able to sense it and thus make changes accordingly.

References

[1] Dien D. Phan,Douglas R. Vogel," Analysis of Promising Beacon Technology for Consumers", Department of Informatics and Quantitative Methods, Faculty of Informatics and Management, University of Hradec Kralove (2016).

[2] https://help.salesforce.com/articleView?id=mc_mp_beacons_overview.htm&r=https%3A%2F%2Fwww.google.com%2F&type=5&sfdcIFrameOrigin=null, last accessed on: 15th October 2020.

[3] Shyam Mohan J S, "A Simple Smart Shopping Application Using Android Based Bluetooth Beacons (IoT)," October , 2017.

[4] http://alerttech.net/5-companies-using-ibeacon-technology-earn-money/, last accessed on: 15th October 2020.

[5] https://altbeacon.github.io/android-beacon-library/, last accessed on: 15th October 2020.

[6] Chen, Guanling, and David Kotz. "A survey of context-aware mobile computing research." (2000).

[7] Schilit, Bill, Norman Adams, and Roy Want. "Context-aware computing applications." 1994 First Workshop on Mobile Computing Systems and Applications. IEEE, 1994.

[8] Jeon, Kang Eun, et al. "Ble beacons for internet of things applications: Survey, challenges, and opportunities." IEEE Internet of Things Journal 5.2 (2018): 811-828.

[9] Faragher, Ramsey, and Robert Harle. "Location fingerprinting with bluetooth low energy beacons." IEEE journal on Selected Areas in Communications 33.11 (2015): 2418-2428.

[10] Liu, Ding-Yu, and Kuei-Shu Hsu. "A study on user behavior analysis of integrate beacon technology into library information services." EURASIA Journal of Mathematics, Science and Technology Education 14.5 (2018): 1987-1997.

A Review of Machine Learning Path Planning Algorithms for Autonomous Underwater Vehicles (AUV) in Internet of Underwater Things (IoUT)

Chinonso Okereke*, Nur Haliza Abdul Wahab and Mohd Murtadha
School of Computing, Faculty of Engineering,
Universiti Teknologi Malaysia, Skudai, Johor, 81310, Malaysia
[e-mail: okereke@graduate.utm.my, nur.haliza@utm.my, murtadha@utm.my]
*Corresponding author: Chinonso Okereke

Abstract

The introduction of the Underwater Internet of Things (UIoT), an extension of the Internet of Things (IoT) underwater has become powerful technology necessary to develop the Smart Oceans. Autonomous Underwater Vehicles (AUVs) play a crucial role in Internet of IoUT technology largely because of their mobility and longer energy storage. However, AUV technologies face major challenges such as path planning problems due to the hostile and dynamic nature of the underwater environment. The path planning problem is about finding an optimal path from the start to the endpoint of the AUV. Machine learning is an approach to tackling this problem. While there are numerous ways to address this challenge, machine learning algorithms are few. This paper provides an overview of the path planning problem and review machine learning path planning algorithms for both single and multiple AUVs and gives directions for future research.

***Keywords*:** Autonomous underwater vehicle, machine learning, path planning, reinforcement learning, neural networks, internet of underwater things, internet of things.

1. Introduction

Internet of Underwater Things (IoUT) is a novel class of the Internet of Things (IoT) that enables Smart interconnected underwater objects. This IoUT allows monitoring and tracking on vast unexplored water areas. Location or object tracking and monitoring in IoT areas such as [1] cannot be applied in underwater things on IoUT. As for that, different methods, approaches need to be explored in IoUT to guarantee fully internet system cover underwater. As about 71% of this earth is covered by the ocean, it is very important to have system support cover in IoUT.

Several technologies developed to fulfill the needs in IoUT such as Autonomous Underwater Vehicle (AUV). AUV plays an increasingly important role in ocean exploration specifically like monitoring, tracking and routing. System applied by the internet specifically in Routing Optimization is also an important role model like in [2] but to handle it underwater it is also such a big challenge. Others like AUV Unmanned Underwater Vehicles (UUV) that are self-propelled and independently operating in six degrees of freedom and can conduct planned missions independently also can be considered great technology applied in IoUT. The other class of UUVs is remotely operated vehicles (ROV)

This research was supported by a UTM Encouragement Research Grant Q.J130000.3851.19J08

which are powered and operated from a station with a cord or remotely [3].

While AUVs have different designs and types such as gliders, hovering and intervention AUVs, the conventional AUV design is usually in torpedo-like shapes because of its advantages such as flexibility, better acceleration abilities, ease of launch and recovery [4]. AUVs are used in ocean exploration as well as mine counter-measures, deep-sea inspections, marine science, security patrols, pipe maintenance, search and rescue in hazardous environments [5].

Due to the importance of AUV technologies, researchers have sought to improve its effectiveness, part of which involves efficiently solving the path planning control problem which is crucial for many applications including data collection, ocean predictions, and monitoring.

According to [6], the path planning problem is defined as calculating a route to a targeted destination which optimizes stated objective functions with the current state of a single AUV or multiple AUVs and ocean environment details. While solving the path planning problem, the characteristics of the vehicle(s) must be maintained.

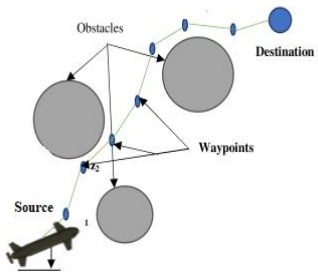


Fig. 1. A typical path planning path for AUV

The objective function to be optimized may be time, energy, or safety depending on the application requirements. Path planning for AUVs has generally been tied to safety conditions. **Fig. 1** shows a typical path planning path for AUV.

1.1 Safety

Safe conditions involve taking a path devoid of obstacles or dangerous areas. A typical vehicle may not have information about the locations of an obstacle. However, as the AUV transverses, through the area, the AUV must have the ability to sense or change their location with time. Other AUVs can also be seen as obstacles in the case of multiple AUVs. The AUV is required to be able to calculate and change its route in real-time. How this is done fulfills the safety objective function [7].

1.2 Energy Consumption

Since AUVs have relatively small battery life, the objective is to keep energy consumption minimal. This can be done by simplifying computational complexities, avoiding obstacles and hazardous areas that can cause unwanted errors, finding the shorter path to destinations, or in some cases reduces the speed of AUV.

1.3 Time Travelled

Time is another objective to be optimized. Increasing the speed of AUV at the expense of energy consumption, avoiding obstacles that cause unnecessary details, and finding short paths of travel are some notable ways of minimizing time spent. It is noted that achieving path objective optimization can be interdependent on one another in more cases.

The prediction of paths along with these specified criteria (e.g., time, energy, data collected, and/or safety) optimized as a whole is therefore labeled path planning [8].

Over the last decade, there has been a lot of research and improvement on path planning in both single and multiple AUV applications. However, most of the development path planning algorithms incorporate little machine learning approaches.

This paper reviews the path planning algorithms with machine learning and seeking to establish the trend and find gaps and areas of possible improvements.

2. Machine Learning

Machine Learning algorithms are generally regarded as computer algorithms that can automatically learn and improve from experience without being explicitly programmed. The three main classes of machine learning algorithms are supervised learning, unsupervised learning, and reinforcement learning.

Supervised learning algorithms train a system, based on examples of each category, to differentiate between different categories or classes of input. Common examples of supervised learning include neural networks, supervised regression support vector machines, adaptive boosting among others. In unsupervised learning, a model is developed by grouping similar unlabelled data. A common type is clustering.

In the case of reinforcement learning, data classification is not needed, rather, the agent learns through trial and error and with the concept of reward and punishment in an environment described by the Markov Decision Process. The most common type of reinforcement learning is Q-learning. The process of reinforcement learning was shown in **Fig. 2**.

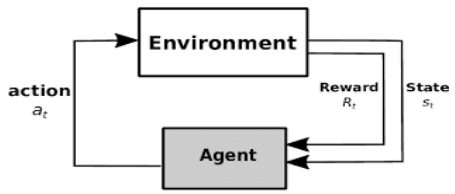


Fig. 2. Reinforcement learning process

2.1 Machine Learning in Path Planning Algorithm

Even though machine learning can assist path planning in other non- planning components of the system like the use of convolutional neural networks (CNN) to automatically segment side-scan sonar (SSS) images in [9] and k-clustering of images for easier computation as shown in [10], the focus this paper is to review the recent path planning algorithms made up of machine learning. The two broad categories of machine learning algorithms used in path planning are neural networks and reinforcement learning.

2.2 Neutral Networks Algorithms

A neural network is a model of a computational network inspired by the structure of an animal brain of biological neurons. It is possible to view the network as a graph of nodes linked by edges. The edges relay activation information from one node to another, analogous to how electrical signals are passed through biological neurons [11].

Table 1 shows the reviewed neural network algorithms in terms of their deployed environment, method, path cost, objective function, availability of obstacle avoidance, and the number of AUVs present

Table 1. Comparison of Supervised learning path planning algorithms

Authors	Type Of Environment	Method	Type Of Path Generated	Obstacle Avoidance	Path Cost	Single/ Multi
[12]	Unpredictable	BINM Neural Network	Optimal	Achieved	Moderate	Single
[13]	Unpredictable	BINN +Velocity Synthesis	Time &Energy Optimal	Achieved	Low	Multi
[14]	Unpredictable	Dynamic BINN (DBINN)	Energy And Time-Optimal	Achieved	Low	Single
[16]	Predictable	Extreme Learning Machine	Time & Energy Optimal	Achieved	Low	Single
[15]	Unpredictable	Glasius BINN	Time And Energy	Achieved	Low	Single
[17]	Unpredictable And Predictable	ANN +Evolutionary Algorithm	Time And Energy Optimal	Achieved	Low	Single

In [12] a topologically organized bio-inspired neurodynamic model based on a sonar map is constructed to represent the dynamic environment and inspire a collision-free path without any prior knowledge of the environment. [13] used an algorithm that combines the Biological Inspired Neurodynamic Model (BINM) and Velocity Synthesis (VS) to produce shorter search paths and thus reduce energy consumption for multiple AUVs compared to the traditional BINM. [14] deals with the shortcomings of BINN such as high computational complexity and long paths for larger environments and bigger obstacles, using

a dynamic BINM.

[15] solves the BINN challenges using a Glasius BINN. In [16] an Extreme learning machine (ELM)s is used to generate an obstacle-free path at a fast speed. [17] combines evolutionary algorithm and artificial neural network to solve multi-objective and multi-stage path planning search operations.

2.3 Reinforcement Learning Algorithms

Reinforcement learning algorithms assume the world is a Markov decision process. An algorithm may try to infer some or all of the MDP by observing the effects of the actions it executes. The resulting estimation of the MDP can then be used to create a policy for future decisions. These types of algorithms are known as model-based methods. In contrast, model-free methods create policies that attempt to maximize reward without modeling the underlying dynamics of the MDP.

2.4 Analysis of Reinforcement Algorithms

Table 2 shows the reviewed reinforcement algorithms in terms of their deployed environment, method, path cost, objective function, availability of obstacle avoidance, and the number of AUVs present.

Han et al [18] use reinforcement learning for the path planning of multiple AUVs alongside an underwater acoustic sensor network for effective monitoring.

In [19], [20], a reinforcement learning algorithm is compared to evolutionary algorithms for path planning. The researchers show that reinforcement learning performs better than the biologically inspired algorithms in higher computational complexity.

Table 2. Comparison of reinforcement learning path planning algorithms

Authors	Type Of Environment	Method	Type Of Path Generated	Obstacle Avoidance	Path Cost	Single/ Multi
[18]	Predictable	Reinforcement Learning	Time And Energy Optimal	Achieved	Moderate	Multi
[21]	Predictable	Reinforcement Learning +Artificial Potential Field	Time-Optimal	Achieved	High	Single
[23]	Unpredictable	Adaptative Dynamic Programming	Optimal	Achieved	High	Single
[22]	Unpredictable	Q-Learning +Path Smoothing Algorithm	Time-Optimal	Achieved	Moderate	Single
[19] [20]	Predictable	Q-Learning And Evolutionary Algorithm)	Optimal	Achieved	Moderate	Single

In [21], reinforcement learning is used for the path planning of intervention AUVs for catching sea urchins in the deep seabed while in [22] reinforcement learning is used in path optimization for a marine vehicle in ocean currents. In [23], the researcher uses a more advanced form of reinforcement learning called adaptive dynamic programming to solve the complex calculations achieving optimal motion control coupled with the least square method.

3. Results Analysis and Recommendations

Based on the reviewed literature from **Table 1**, it is seen there has been more focus on the application of neural networks in single AUVs (83%) than in multiple AUV systems (17%) and the type of environment for the application of neural networks are mainly unpredictable (83%). While all reviewed literature achieved obstacle avoidance, most of them were also both time and

energy optimal (83%) in the path generated with every low path cost. Also

For **Table 2**, it is seen also that more focus has been given to single AUV applications (80%) than multiple AUV applications (20%). Just like other machine learning algorithms, obstacle avoidance is achieved, however, few reinforcement learning algorithms (20%) achieve target energy optimization and the path cost of the reinforcement algorithms ranges from moderate to high. It is also noted that the type of environment that applies reinforcement algorithms are 60 %predictable and 40% unpredictable based on the reviewed literature. With this analysis, the following recommendation is given for future study:

There is a need for more research into the use of machine learning in the path planning of multiple autonomous underwater vehicles (cooperative path planning).

Also, from the reviewed literature it shows that reinforcement learning is more energy-intensive than neural network and other machine learning approaches in path planning. There is therefore needed to research more energy-efficient methods of path planning with reinforcement learning.

4. Conclusions

This paper reviews machine learning approaches to path planning. The main machine learning algorithms used in path planning include neural networks and reinforcement learning. While there has been a good amount of research into machine learning approaches in path planning, there are still several milestones to be achieved like energy efficiency with reinforcement learning and effective machine learning approaches in cooperative path planning.

References

[1] Nur Haliza Abdul Wahab, Sharifah H.S Ariffin, Liza Abdul Latiff, Sarerusaenye Ismail, "Indoor Location Assistant by Integrated Localized Routing in Proxy Mobile", *Journal of Advanced Research in Dynamical and Control Systems*, 11(10 Special Issues), pp. 1108-1115, October 2019.

[2] Nur Haliza Abdul Wahab, Latif, L.A, Ariffin, S.H.S, Ghazali N. E, "Route optimization via RSSI APPS in indoor proxy mobile IPv6 test-bed", *Research Journal of Applied Sciences, Engineering and Technology*, 2014, 8(8), pp. 942-951.

[3] M. Panda, B. Das, B. Subudhi, and B. B. Pati, "A Comprehensive Review of Path Planning Algorithms for Autonomous Underwater Vehicles," Int. J. Autom. Comput., vol. 17, no. 3, pp. 321–352, 2020, doi: 10.1007/s11633-019-1204-9.

[4] K. Alam, T. Ray, and S. G. Anavatti, "A brief taxonomy of autonomous underwater vehicle design literature," Ocean Eng., vol. 88, pp. 627–630, 2014, doi: 10.1016/j.oceaneng.2014.04.027.

[5] N. Kumar and M. Rani, "An efficient hybrid approach for trajectory tracking control of autonomous underwater vehicles," Appl. Ocean Res., vol. 95, no. January, p. 102053, Feb. 2020, doi: 10.1016/j.apor.2020.102053.

[6] Y. Kuwata, T. Schouwenaars, A. Richards, and J. How, "Robust constrained receding horizon control for trajectory planning," Collect. Tech. Pap. - AIAA Guid. Navig. Control Conf., vol. 3, no. August, pp. 2375–2386, 2005, doi: 10.2514/6.2005-6079.

[7] Z. Zeng, L. Lian, K. Sammut, F. He, Y. Tang, and A. Lammas, "A survey on path planning for persistent autonomy of autonomous underwater vehicles," Ocean Eng., vol. 110, pp. 303–313, 2015, doi: 10.1016/j.oceaneng.2015.10.007.

[8] P. F. J. Lermusiaux et al., "A future for intelligent autonomous ocean observing systems," J. Mar. Res., vol. 75, no. 6, pp. 765–813, 2017, doi: 10.1357/002224017823524035.

[9] P. Liu and Y. Song, "Segmentation of sonar imagery using convolutional neural networks and Markov random field," Multidimens. Syst. Signal Process., vol. 31, no. 1, pp. 21–47, 2020, doi: 10.1007/s11045-019-00652-9.

[10] R. J. Wai and A. S. Prasetia, "Adaptive Neural Network Control and Optimal Path Planning of UAV Surveillance System with Energy Consumption Prediction," IEEE Access, vol. 7, pp. 126137–126153, 2019, doi: 10.1109/ACCESS.2019.2938273.

[11] M. W. Otte, "A Survey of Machine Learning Approaches to Robotic Path-Planning," Int. J. Rob. Res., vol. 5, no. 1, pp. 90–98, 2008, doi: 10.1109/ICALIP.2016.7846622.

[12] J. Ni, L. Wu, P. Shi, and S. X. Yang, "A dynamic bioinspired neural network based real-time path planning method for autonomous underwater vehicles," Comput. Intell. Neurosci., vol. 2017, 2017, doi: 10.1155/2017/9269742.

[13] D. Zhu, C. Tian, B. Sun, and C. Luo, "Complete Coverage Path Planning of Autonomous Underwater Vehicle Based on GBNN Algorithm," J. Intell. Robot. Syst. Theory Appl., vol. 94, no. 1, pp. 237–249, 2019, doi: 10.1007/s10846-018-0787-7.

[14] D. Dong, B. He, Y. Liu, R. Nian, and T. Yan, "A novel path planning method based on extreme learning machine for autonomous underwater vehicle," Feb. 2016, doi: 10.23919/oceans.2015.7401951.

[15] N. Abreu and A. Matos, "Case-based replaiming of search missions using AUVs," in OCEANS 2017 - Aberdeen, Oct. 2017, vol. 2017-Octob, pp. 1–10, doi: 10.1109/OCEANSE.2017.8084990.

[16] M. Yan, F. Gao, X. Qin, and D. Zhu, "Sonar-based local path planning for an AUV in large-scale underwater environments," Indian J. Geo-Marine Sci., vol. 46, no. 12, pp. 2527–2535, 2017.

[17] X. Cao and D. Zhu, "Multi-AUV Underwater Cooperative Search Algorithm based on Biological Inspired Neurodynamics Model and Velocity Synthesis," J. Navig., vol. 68, no. 6, pp. 1075–1087, 2015, doi: 10.1017/S0373463315000351.

[18] G. Han, Z. Tang, Y. He, J. Jiang, and J. A. Ansere, "District Partition-Based Data Collection Algorithm with Event Dynamic Competition in Underwater Acoustic Sensor Networks," IEEE Trans. Ind. Informatics, vol. 15, no. 10, pp. 5755–5764, Oct. 2019, doi: 10.1109/TII.2019.2912320.

[19] Y. Noguchi and T. Maki, "Path Planning Method Based on Artificial Potential Field and Reinforcement Learning for Intervention AUVs," Apr. 2019, doi: 10.1109/UT.2019.8734314.

[20] S. Vibhute, "Adaptive Dynamic Programming Based Motion Control of Autonomous Underwater Vehicles," in 2018 5th International Conference on Control, Decision and Information Technologies, CoDIT 2018, Jun. 2018, pp. 966–971, doi: 10.1109/CoDIT.2018.8394934.

[21] B. Yoo and J. Kim, "Path optimization for marine vehicles in ocean currents using reinforcement learning," J. Mar. Sci. Technol., vol. 21, no. 2, pp. 334–343, 2016, doi: 10.1007/s00773-015-0355-9.

[22] U. Gautam and M. Ramanathan, "Simulation for path planning of SLOCUM glider in near-bottom ocean currents using heuristic algorithms and Q-learning," *Def. Sci. J.,* vol. 65, no. 3, pp. 220–225, 2015, doi: 10.14429/dsj.65.7855.

[23] U. Gautam, R. Malmathanraj, and C. Srivastav, "Simulation for path planning of autonomous underwater vehicle using Flower Pollination Algorithm, Genetic Algorithm and Q-Learning," Apr. 2015, doi: 10.1109/CCIP.2015.7100710.

Forecasting Hourly Electricity Trading Volume based on Weather Forecast Data Using Deep Learning

Junmo An[1], Hoonseok Park[1,2], Hanee Ryu[3] and Jae-Yoon Jung[1,2]
[1]Department of Industrial and Management Systems Engineering, Kyung Hee University, South Korea
[2]Department of Big Data Analytics, Kyung Hee University, South Korea
[3]Business and Innovation Department, Hyundai Research Institute, South Korea
[e-mail: ajm9306@khu.ac.kr, hoonseok@khu.ac.kr, hnryu@hri.co.kr, jyjung@khu.ac.kr]
*Corresponding author: Jae-Yoon Jung

Abstract

The power trading volume is an important indicator that contains information on both the amount of power demanded and power generation to analyze the cost-based generation pool (CBP). With the increase in demand and generation of renewable energy, the power trading volume is directly or indirectly affected by weather conditions. In this study, we propose a method of forecasting the power trading volumes from the current weather conditions using a deep learning model based on the stacked long short-term memory (LSTM). First, the weather observation data and the past power trading volume data were combined and preprocessed. Then, the LSTM-based deep learning model, called WE-Net, was designed and compared with existing statistical models and machine learning models. The hourly electricity trading volume and weather observation data from 2015 to 2017 in Jeju, Korea were used to evaluate the proposed method. As a result of this study, it was confirmed that the proposed WE-Net showed higher performance than conventional forecasting methods including statistical models and machine learning techniques.

***Keywords*:** AMGO (Adjusted Metered Generation Output), electricity trading volume, weather variable, deep learning, time-series prediction

1. Introduction

Electricity trading is made by determining the market price according to demand forecast and generation cost calculation for each generation source. At this time, the electricity trading volume includes information on both the amount of demand and generation, and becomes a key index for analyzing CBP [1] . This means that the amount of electricity traded in the electricity market functions as a signal for an important electricity supply and demand situation.

With the increase of renewable energy generation, studies on predicting a volume of electricity generation and supply using weather variables have been proposed [2-4]. However, there is no research on direct forecasting using weather variables.

In this study, we propose a stacked long short-term memory (Stacked-LSTM) model of forecasting electricity trading volume based on weather variables. Adjusted metered generation output (AMGO) is considered the trading volume. The definition of significant time variables and pre-processing of weather

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2017H1D8A2031138).

variables were included.

The rest of this paper is organized as follows. In Section 2, the structure of the proposed model, WE-Net (Weather based Energy trading prediction Network), is explained. In Section 3, the model is evaluated using past AMGO and weather observation data in the Jeju. Finally, Section 4 concludes this paper with conclusions and future research.

2. Method

In this study, we present a deep learning model, called WE-Net, for forecasting hourly AMGO using weather data. WE-Net resembles a stacked LSTM model [5]. Accordingly, WE-Net receives a sequence of past weather data and trading volume information for a certain time as input, and outputs the electricity trading volume by time for a certain time after the input time as a prediction. In this study, we learn WE-Net, which receives weather information for 24 hours and historical AMGO data as input and predicts the hourly AMGO for 6 hours thereafter. The architecture of the proposed WE-Net is shown in **Fig. 1**.

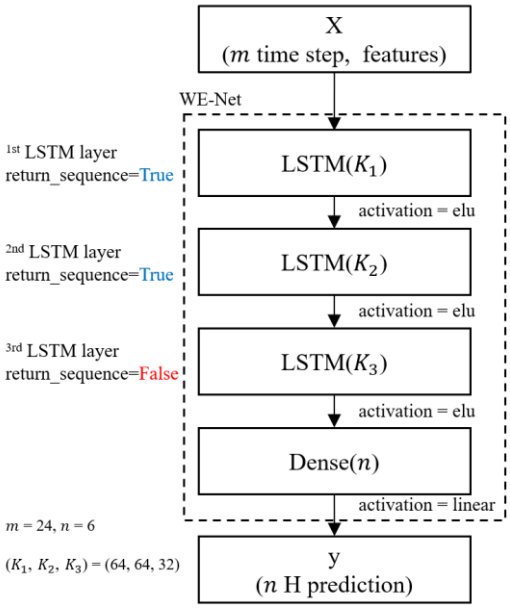


Fig. 1. WE-Net architecture

The input shape of WE-Net is (m time step, feature). The time step is a parameter for how many hours of data will be used by WE-Net to predict the AMGO for after n hours. In this experiment, 24 for m time step parameter and 6 for prediction length n. The feature parameter is the dimension of a variable, and in this experiment, the data categorized through one-hot encoding has a feature parameter value of 47. The input X data goes through three LSTM layers, and the return_sequences parameter is set to True in the first two layers, so that the output form of each layer returns a sequence equal to the size of the timestep. Afterwards, the last LSTM layer sets the return_sequences parameter to a False value to finally return data with compressed information. In order to predict the AMGO for the next 6 hours, the Dense layer is composed of 6 nodes, and the result is returned by compressing the previous temporal information.

3. Experiments

3.1 Data Preprocessing

The data for the research is used by combining the data on the amount of electricity traded by time (AMGO) and weather observation data for about 3 years in the Jeju. The shape of the variable consisted of numerical variables such as AMGO, temperature, humidity, and categorical variables such as day of the week and holidays.

In the data pre-processing step, weather observation data processing, missing value processing, data combining, normalization, and time-sliding windowing are conducted to construct a dataset to be used in the AMGO prediction modeling considering weather information over time. For two missing values of AMGO, it was interpolated as the average value of their two adjacent AMGO. Finally, AMGO and weather observation data were combined into a single dataset based on time.

Table 1. Learning data variables

Type	Variable name
Weather variable	Temperature (˚C), Humidity (%), Wind speed (m/s), Wind direction (rad), Air pressure (hPa), Sunshine duration (m), Solar insolation (kWh), Visuality (10m)
Time variable	Hour, Day of the week, Season, Holiday, Number of continuous holiday, n_th holidays

The data used in this paper is a multivariate dataset with weather and time variables, and a model returns a prediction of a certain period (6 hours) through input for a certain period (24 hours). To train this model, the preprocessed dataset was reconstructed using the time window sliding technique of multi-step forecasting for multivariate datasets. Through this process, the shape of the dataset is configured so that the model to be trained uses information over a certain period of time for learning.

In this study, the training dataset for the experiment and the test set for verifying the performance were divided according to the period. The training dataset is about 2 years and 6 months (130 weeks) from 2015-01-01 to 2017-06-30, the test dataset consisted of about 2 months (8 weeks) from 2017-07-01 to 2017-08-25.

3.2 Model Training and Evaluation

WE-Net was trained using the processed data, and the performance was compared with existing statistical time series and machine learning techniques. As the univariate time-series model, ETS (error, trend, seasonal) was used, and as machine learning models, ensemble algorithms such as random forest (RF), extreme gradient boosting (XGBoost), light gradient boosting machine (LGBM), and CatBoost, which generally exhibit high performance, were used as shown in **Table 2**.

Table 2. Model evaluation metrics

Type	Model	R-squared	RMSE	MAPE
Time-series method	ETS	0.740	45.728	32.732
ML methods	RF	0.746	45.338	34.722
	XGBoost	0.737	45.107	35.359
	LGBM	0.739	45.935	34.940
	CatBoost	0.745	45.390	35.039
Proposed method	WE-Net	**0.775**	**42.663**	**32.316**

The proposed WE-Net algorithm showed best performance in terms of R-squared, root mean squared error (RMSE), and mean absolute percentage error (MAPE) compared to the existing techniques.

4. Conclusions and Future Work

In this study, we proposed a Stacked-LSTM based WE-Net model that effectively forecasts the electricity trading volume using weather variables. WE-Net outperformed the existing methods such as statistical time series and machine learning techniques. The proposed method was verified with AMGO and weather observation data in Jeju, South Korea.

The 24-6 hour input/output format applied in this study may have limitations in being directly used to predict the actual power trading volume. WE-Net is expected to increase the relative performance when predicting longer-term data. In order to increase the relative performance and usability of the proposed model, it is necessary to consider output over 24 hours.

References

[1] S.-M. Han and B.-H. H. Kim, “A Study on the Generation Expansion Planning System Under the Cost Based Pool”, The Transactions of The Korean Institute of Electrical Engineers, Vol. 58, No. 5, pp. 918-922, 2009.

[2] J. R. Andrade and R. J. Bessa, “Improving renewable energy forecasting with a grid of numerical weather predictions,” IEEE Transactions on Sustainable Energy, Vol. 8, No. 4, pp. 1571-1580, 2017.

[3] K. Park, J. Song, and E. Hwang, “Hybrid Load Forecasting Model based on Atypical Residual of Meteorological Information”, KIISE Transactions on Computing Practices, Vol. 25, No. 1, pp. 52-57, 2019.

[4] R. Corizzo, M. Ceci, H. Fanaee-T, and J. Gama, “Multi-aspect renewable energy forecasting,”, Information Sciences, Vol. 546, pp. 701-722, 2020.

[5] F. A. Gers, J. Schmidhuber, and F. Cummins, “Learning to forget: continual prediction with LSTM”, 9th International Conference on Artificial Neural Networks, pp. 850-855, 1999.

Analyzing Senior content trend in Social Media Service

Yoosin Kim*, Doohong Min*, Yeonjin Ju*, Heejoo Kang*
Alticast Inc. Seoul, South Korea
[e-mail: yoosin@alticast.com]
*Corresponding author: Yoosin Kim

Abstract

Currently, the use of social media and YouTube is exploding among middle-aged and elderly people, and the demand for creating and distributing digital content directly from these elderly people is also rapidly increasing. Therefore, in this study, we would like to research and analyze major topics by collecting and analyzing various contents (SNS, online media contents) that seniors actually produce and consume, and big data of their empathy (reaction like good and bad, comments etc.). To this end, AI technologies such as deep learning-based object detection, which recommend popular genres and functions by extracting meta information of digital video content, classifying, and learning content properties, will be applied in the future. The results of this study are expected to be basic analysis data that can provide results of trend analytics that help in content creation to senior creators.

***Keywords*:** Big Data, Trend Analytics, LDA, Topic modeling, Senior creator

1. Introduction

Changes in the silver industry are taking place rapidly due to aging. Currently, the proportion of older people is constantly increasing. In addition, the cultural needs of seniors and the related silver industry are also expanding. The senior generation is also actively consuming online platforms, and develops products and contents directly. As of December 2018, the number of YouTube users who is over 50 in Korea was 9.48 million, up 24% from a year ago (7.62 million), and senior YouTube users are said to watch an average of 922 minutes a month. There is a trend of increasing number of seniors who challenge creators directly, not just watching. To this end, this study analyzes current content trends to meet the growing needs of seniors. It will be helpful for first-time senior creators to understand trends of creation and also help for improving content quality and direction of production.

2. Related Work

3.

Trend analysis and topic modeling analysis are in progress in various fields. Il-seop Lee et al [1] developed the study was conducted to find out trends on the Internet through SNS data and spatial analysis. In K-pop trend analysis using Yang Geun-woo's[2] melon chart data, using keyword data extracted through morpheme analysis of song titles and lyrics, a network analysis technique was applied to identify the K-pop trend and its success factors. And Kim Young-ah et al. [3] proposed to Using the big data of a news article, an analysis of the trend of elderly jobs in 2010-2019 was conducted. Using news data collected through web crawling and Naver Data Lab search terms, various data such as word cloud, network graph, and ego network were visualized, and also what keywords appeared and the trends were examined in relation to senior's jobs. Also, text mining for topic modeling is also being actively researched,

This research is supported by Ministry of Culture, Sports and Tourism and Korea Creative Content Agency(Project Number: R2020040068)

and text frequency analysis was conducted to find out the overall trend of consumer reaction toward overseas direct purchase[4]. In addition, in the research of Park Geon-cheol et al. [5], topic modeling using the LDA technique was performed to understand the trend of smart city research.

4. Analytics

5.

Senior content trend analysis was conducted through data collection, preprocessing, and analysis. In the data collection stage, we investigated the data collection channel related to the content of senior creators and reviewed the data collection method. In the data pre-processing stage, some information and stop words that were less related to the analysis were removed, and the analysis was carried out by topic modeling using LDA.

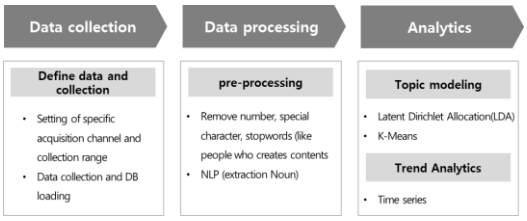


Fig. 1. Process for data analytics

3.1 Data Collection

For data collection, we selected two SNS platforms(Youtube & Naver) that are most often used by seniors. 16 video-based SNS creators and one text-based channel were selected.

3.2 Data Preprocessing

Data was pre-processed before data analysis. First, numbers and special characters were removed. Second, a stopword dictionary created through prior research was created, and the stopwords were removed using this. Third, noise that may affect the analysis result was removed. The nouns were extracted using the Python KoNLPY library. Through the above steps, data that can be used for data analysis was constructed.

3.3 Data Analytics

Through the collected 5,901 data, we tried to understand the overall preferences and interests of seniors. For this purpose, keywords with too few occurrences were excluded. Also, common keywords that appear in the document are excluded because it is difficult to say that they have meanings.

4. Analysis

4.1 Topic Modeling

Based on the topic analysis results, the interests of the seniors could be confirmed. The 8 topics derived can be found in Table.1. As can be seen in Figure.3, the proportion of "Topic-5 Finance/Retirements (17.3%)" was the highest. "Topic-2 Finance/Retirements (16.4%)", "Topic-1 Mukbang (13.2%)", "Topic-4 Politics (12.9%)", "Topic-8 Agriculture (11.3%)", "Topic-3 Topics were delivered in the order of Daily Life (10.9%)", "Topic-7 Health/Welfare (10.8%)", and "Topic-6 Community (7.2%)". The interest in life and politics after retirement was highest. And it is noteworthy that seniors are also interested in materials such as “mukbang”, which seemed to be limited to young people.

In addition to these topics, we have also found that seniors run channels on topics related to their specialty. For example, senior creator "Milanonna" has worked in the fashion world for a long time, and currently runs fashion-related channels.

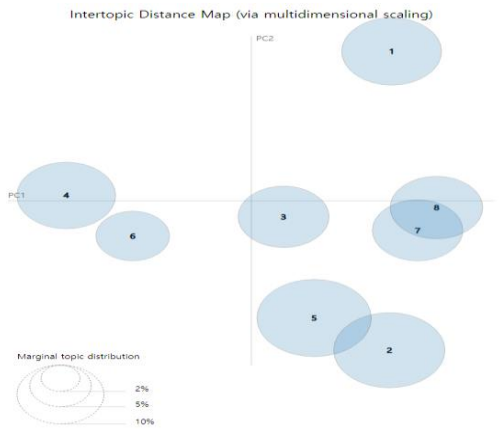


Fig. 2. Intertopic Distance Map

4.2 Trend Analysis

The number of documents per topic was analyzed on a weekly basis in order to grasp the change of the senior's interest over time. Documents from July 2020 to September 2020

were targeted to confirm the latest time series trends. Overall, interest in "Topic-2 Finance/Retirements", "Topic-8 Agriculture" and "Topic-7 Health/Welfare" was high. Recently, it was confirmed that the proportion of "Topic-3 Daily Life" has increased slightly.

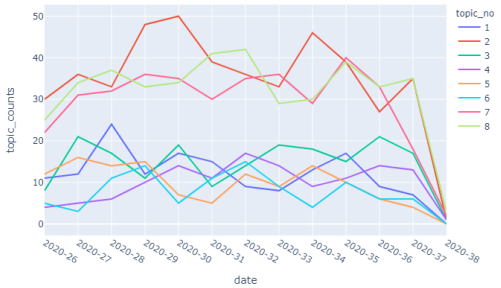


Fig. 3. Trend in Topic Change by week (2020-07 ~ 2020-09)

5. Conclusions

Through this study, topic analysis of SNS contents produced and consumed by seniors was conducted. When looking at the analysis results, the proportion of topics such as retirement, return to farming, and health was high. Based on this, it was found that seniors are highly interested in retirement and life after retirement. During this study, there were some difficulties in preprocessing Hangul data. Because, it is necessary to process stop words and analyze morphemes by considering the context and synonyms of Korean. This process took a relatively long time. In order to solve this problem, we need to continuously manage stop words.

In order to use the research results better, it is necessary to additionally collect various SNS channel data and SNS data of the current channel. This is because it is difficult to show the trends/topics of the entire senior media content with the currently collected data. Also, in addition to topic analysis using LDA, it is necessary to use the other anaysis technique like Word2Vec or Network Analysis for each topic. And based on this, it can enhance the analysis results. And based on this, it is necessary to enhance the analysis results.

In the future, through a deep learning AI-based service that extracts meta-information of digital video content, classifies and learns content properties, and recommends popular genres, it will be possible to provide customized information that is useful for content creation for senior creators. This study can be used as basic analysis data for this.

References

[1] Il-sup Lee, Kyeong-kyu Kim, and Ae-ri Lee, "A Big Data Analysis Methodology for Examining Emerging Trend Zones Identified by SNS Users: Focusing on the Spatial Analysis Using Instagram Data", *Information Systems Review,2018* Vol. 20, pp. 63-85, 2018.

[2] Yang, Kun-Woo, " K-POP Trend Analysis Using Melon.com Chart Data," *The e-Business Studies,* Vol. 18, pp. 3-19,2017.

[3] Kim Young-a, Lee Hye-jin, Cha Su-min, Park Yoon-he., " The trend of the elderly jobs using news article Big date: 2010-2019.", *The Journal of Korea Aging Friendly Industry Association*, Vol. 12, pp.13-20,2020.

[4] Jong Youn Rha, Jin Myong Lee. "Exploring Consumer Responses to the Cross-Border E-Commerce using Text Mining". *JOURNAL OF CONSUMER STUDIES*, Vol. 26, pp.93-124, 2015.

[5] keonchul Park, Chi Hyung Lee, “A Study on the Reserch Trends for Smart City using Topic Modeling”, *JICS* Vol 20(3), pp. 119-128, 2019.

Reusable Patient Navigation System Architecture using Component Based Approach

Muhammad Syafiq A. Manan, Dayang N.A. Jawawi, Shahliza Abd Halim
School of Computing, Faculty of Engineering
University Teknologi Malaysia Johor, Malaysia
E-mail : msyafiq364@graduate.utm.my, dayang@utm.my, shahliza@utm.my

Ushananthiny Suvelayutnan, Mohamad Noor Hakim Mohamad
Hospital Sultanah Aminah, Johor, Malaysia
E-mail : ushananthiny@moh.gov.my, m.hakim@moh.gov.my

Abstract

Patient Navigation Program (PNP) had started being implement since 1990 which the initial concept introduced by Horland P. Freeman. This program is aimed to reduce the barrier access the treatment and diagnosis of the cancer patients especially for the poor, total disability, and the conditioning. Nowadays, the idea of Patient Navigation System (PNS) starts to develop and most of the healthcare centers are moving toward to more computerized system. In Malaysia, Hospital Sultanah Aminah (HSA), Johor already implemented PNP to navigate their patient for rehabilitation unit, but they are still using the paper documentation to execute the program. Therefore, a PNS for Rehabilitation of Paralysis Patient as a platform had been developed to ensure that patient data organized more systematically. The current architecture of this software was using three-layer architecture by applying the component-based approach. However, this approach provides small scale of reusable features. Since the hospital process usually complex and changes time by time. Thus, the development of PNS architecture needs to be reusable for varies other functionality changes. Currently, the increasing use of software product line (SPL) approach which offer a huge advantage in software development concerning the cost, time, and effort reduction in large scale of reusable features. This paper proposed to adapt the component-based in SPL to develop the reusable PNS architecture by merge two of this approach advantages. By adapting both of these approach to develop the reusable PNS architecture could result on a decent level of reusable software architecture for the future development of PNS.

***Keywords*:** Patient Navigation System; Rehabilitation; Component-Based; Software Product Line

1. Introduction

Patients Navigation Program (PNP) are aimed to assist and coordinate the cancer patients after identified the abnormal cell detected on the body according to the cancer process and procedure at the beginning of its history [1]. The initial concept of this program is implemented from the idea and detail concept that had been done by Harold P. Freeman in Harlem, New York in 1990. The studies from Patient Navigation Research Program show that this program can reduce the amount of time of the treatment for the long- term diagnosis patients.

Presently, the patient navigation is not only emphasized the cancer patients but also had also been using for rehabilitation patients etc. This treatment also needs a very long-term diagnosis

on the patients. To ensure that the patients can get the specific treatment properly especially involving the repetition treatment and diagnosis. This program had widespread used in varies domain of diseases, due to it beneficial that the patients and healthcare will gain from it.

Information and Communication Technology (ICT) had influence most of the daily work of human activities. It also influences the way healthcare to manage their service. The process of the hospital section usually complex and changes time by time. Thus, the development of PNS architecture needs to be reusable for varies other healthcare facilities that shared the same domain under PNS scope. This could reduce the time, cost and effort for the future development of the PNS.

The aim of this paper is to produce the reusable PNS architecture by adapting the components based and software product line approach. On this project, the existing system of PNS using component based approach will be adapt with the SPL to produce the reusable architecture for this system.

2. Related Work & Background

In this section, the related work of the project development will be discuss more details.

2.1 Patient Navigation Program

PNP had being implement in varies of the long-term treatment disease around the world. Patient Navigation Program is the mechanism which navigate the patients that undergoes long-term treatment especially provide healthcare service to cancer patients by abolish the barriers and manage patient's engagement to the cancer care services to seek treatment appointment time properly.

In Malaysia, this program had been implemented by Hospital Tengku Ampuan Rahimah (HTAR) associated with Cancer Research Malaysia (CRM) which mainly aim to ensure the continuity and survivorship of the patients by surpass the barriers of communication between patients and healthcare. There is a lot of initiative that had been taken to ensure that this patient navigation program keep on its good track by the Ministry of Health Malaysia (MoH). **Fig. 2** shows the Expend of Patient Navigation Model [2].

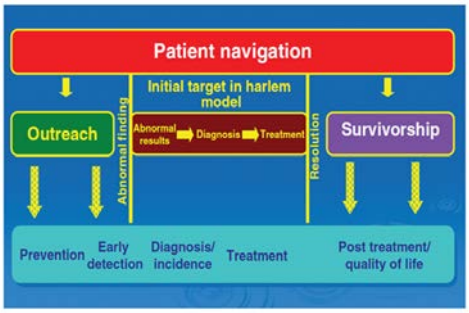


Fig. 2. Expand Navigation Model by Freeman (2011)

Little research had been done on the reuseable architecture PNP system for healthcare around Malaysia. Currently, SPL is the approach that had been used to develop HTAR PNP core asset for breast cancer case study. This approach systematically encompass all aspects of the software development process from the requirement elicitation until the implementation of the system while promising the maximizes reusability of the core assets in respective domain.

Next, HSA had also implemented this program to their rehabilitation unit to keep track all the patient data. The elicitation of functionality and non-functionality requirement had being done and analysed. **Fig. 3** show the current workflow of PNP that at HSA.

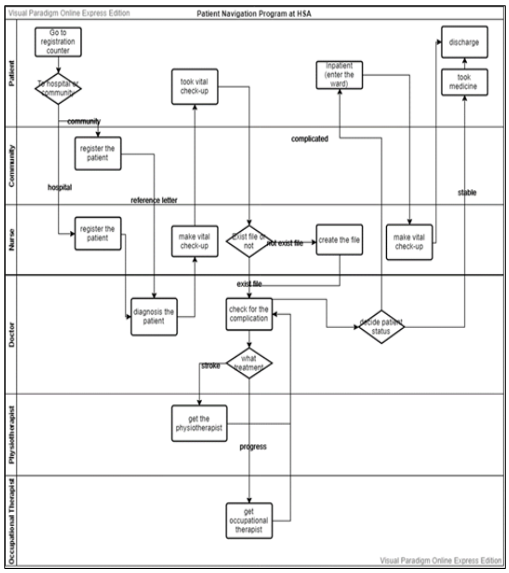


Fig. 3. Existing Workflow for Patient Navigation Program at HSA

2.2 Software Reuse Approach

C.Szyperski defined the software components is an executable unit of the code that offer the specified set of interfaces [3]. On this study, the component based development emphasized on the separation of concern (SOC) for all the aspects in the functionality of the system requirements. This approach considers being a reused based approch with offer loose coupling in the independent components.

By applying this approach, software systems developments can be developed by selecting appropriate components and assembling it with suitable software architectures [4]. Each of the system development can be significantly reduced the time,cost and effort to deploy the system.

Furthermore, software product line engineering (SPLE) is one of the reuse approaches. SPLE is a set of applications that can be generalized around a common architecture that could be adapted with the different stakeholders. SPLE is about extracting the commonalities of the domain and systematically manage the variabilities [5]. The main part of the system in this approch is the core assets that become fundamental for the system developments.

When this core assets had being set and established, the development of the product line of the system will become more secure, faster and less costs. By implementing this approach it give significant impact on the development of the system, especially the development was emphasized on same domain of the product line. This could give the advantage to the developer to shorter the time to developed the system.

SPL development process can be split into two part which is domain engineering (DE) and Application engineering (AE). For domain engineering, it will focus on defined the commonalities and variabilities of the SPL which later produced the reused core asset of the system. While for application engineering, is focus on to realized the core assets from the DE platform of the specific product line by consider the commonalities and variabilities of the SPL [6]. **Fig. 4** shows the SPL processes illustrated by Clements and Northrop [7].

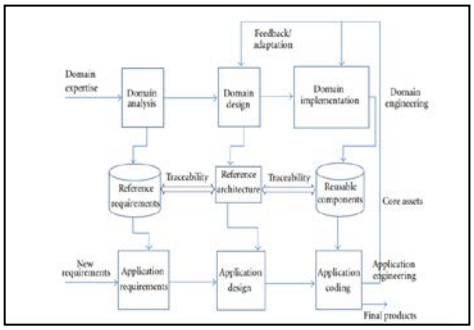


Fig. 4. SPLE processes by Clements and Northrop (2002)

Nowadays, various technologies adapted the software reuse approach are available such as DevOps. DevOps aims at unifying software development and operations to improve products and deliver value to customers [8]. This enterprise capability provides continuous software delivery that allows the clients to seize the industrial opportunities. Currently, IBM DevOps Solution had offer built on an open standards platform with the various set of services.

Unlike SPL and component-based approach, DevOps had composed all of the elements in software development into a platform as a services. This interesting features give the advantages to DevOps for more advanced development of the software implementations.

3. The Implementation of the Reusable Approach

In this section, the focus are more PNS system architecture including the components based and SPL approach specific on the PNP domain case study. This implementation adapt the SPL domain engineering phase development in compoenet-based approach.

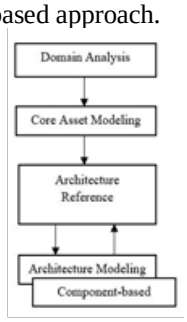


Fig. 5. Domain Engineering for this reused approach

3.1 Software Product Line in PNS

By using SPLE, a SPL reused core assets was created to facilitate the for new development of software products in PNP domain. With various of use for PNP in other long-term treatment disease. Thus, this PNS with reusable architecture should be produced as a framework for future development of the system which share the same domain.

The SPL process start from domain analysis applying Feature-Oriented Reuse Method (FORM) approach. From here, the functional requirement, non-functional requirements, commonalities and variabilities of the PNP domain can be defined. The variability management is the importance aspect that influence the degree of success of SPL [9].

Then, the variability management technique from Brisaboa et al. [10] and variability types from Apel et al. [11] adapted. The PNP features then being classified into capability, operational, domain technology and implementation technique category to produce traceability table. Next, the features in traceability tables being mapping to reference architecture.

3.2 The Component-Based Approach in PNS

After completing the domain analysis for the PNP, the next step of the development will focus on architecture modeling. In this process, the architecture reference will become the source of reference to produce detail design of the system. For example, in component diagram, each of the subsystems will be split into specific units of task execution.

The functionality is break down into separate components that overlaps as minimum as possible by considering the principle of separation of concerns (SOC). Each of the components can independently be executed and communicated through interface. Figure 6 shows the component diagram that had been developed based on the system requirements.

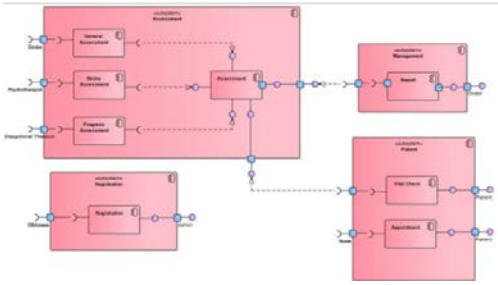


Fig. 6. Component Diagram PNS for HSA rehabilitation unit

3.3 Layer Architecture Pattern in PNS

For the architecture pattern, Layered Architecture Model selected to develop PNS include the View Layer, Business Layer, and Database Layer. During the development of this system, this architecture is used to planning overall system. For View Layer is will consist of the interface, business Layer will consist of the activities that take place in the subsystem, and the database layer is where the data is store and retrieve from the system.

The concern here is to be organized a system into separated layer by considering the separation of concern (SOC) principle. The layer is arranged into successive layers in such a way each of the respective layers responsible for its own layer or below [12]. Figure 7 shows the layered architecture for PNS.

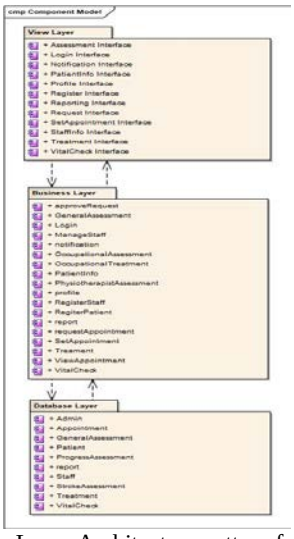


Fig. 7. Layer Architecture pattern for PNS

4. Result

In this section, the result and discussion based on the implementation of the component-based and SPL approach to develop the PNS will be emphasized.

4.1 Adapt the Component Based to SPL approach

In the previous section, we had discussed the implementation of component-based and SPL to produced reuseable PNS architecture. In this paper, for component-based development it more focuses on the details design of the system by applying the separation of concern technique while for SPL as general it will offer architecture support PNP domain to be reuse by produce a PNP system product line.

The component-based approach can be adapt during the domain engineering activities take place start from the domain design and domain implementation which is more emphasized to the details design of the system. Moreover, the architecture reference (core assets) from HTAR PNS development can be extended to include specific requirements for HSA PNS such as functionality commonalities and the architecture pattern used.

Both of approaches offer reusable features which bring both approach together will result on the reduction of time, cost and effort to deliver the product to the stakeholders. Since both of this approach provides good reuseable features for software development. Besides, component-based development also can overcome the lack of maturity of SPL approach by offering the various technologies of development [13].

Since our approach needs more time to well define, this paper also presented the outlines for PNS development proposed this approach which is focus on the component-based development and SPL approach.

5. Conclusion

This paper presented the reusable PNS architecture by adapting component based development in SPL approach which result on high level of reusable architectures. Component-based software development is more focus on small scale of reuseable features while for SPL approach it promote the reusable features for a huge scale software.

This paper also presents the case study of PNS at HSA that applying the component-based system development which consists of the basic components for PNS that had been done in HSA. Then, adapting SPL approach could offer the general framework to developed the PNS in future development. The details of this PNS reference architecture could be improved time by time based on the requirements.

Acknowledgment

Upon completion this research, we would like to express our gratitude to many parties that was involved in this project.

The authors fully acknowledge Universiti Teknologi Malaysia for UTM-TDR Grant Vot No. 06G23, and Ministry of Higher Education (MOHE) for FRGS Grant Vot No. 5F117, which have it possible for this research effort. The authors would also like to express their warmest appreciation for continuous support and input to the members of Embedded Real-Time Software Engineering Laboratory and Software Engineering Research Group.

References

[1] Freeman, H.P., “Patient navigation: a community centered approach to reducing cancer mortality”. Journal of Cancer Education, 21, 2006.

[2] Freeman, H.P. and Rodriguez, R.L., History and principles of patient navigation. Cancer, 117(S15), pp.3537-3540, 2011.

[3] C. Szyperski. “Component Software”. Addison-Wesley, Harlow, England, 1998.

[4] Xia Cai, Michael R. Lyu, Kam-Fai Wong, Roy Ko, “Component-Based Software Engineering:Technologies,Development Framework, and Quality Assurance Schemes”, 2000.

[5] Gomaa, H. “Designing software product lines with UML” (pp. 160-216). IEEE, 2005.

[6] Bosch, J., Florijn, G., Greefhorst, D., Kuusela, J., Obbink, J.H. and Pohl, K., "Variability issues in software product lines. In International Workshop on Software Product-Family Engineering" (pp. 13-21). Springer, Berlin, Heidelberg, 2001.

[7] Clements, P. and Northrop, L., "Software product lines: practices and patterns:" (Vol. 3). Reading: Addison-Wesley, 2002.

[8] Jiménez M., Rivera L.F., Villegas N.M., Tamura G., Müller H.A., Gallego P., "DevOps' Shift-Left in Practice: An Industrial Case of Application", 2019.

[9] L. Chen, M.A. Babar, A. Nour, "Variability Management in Software Product Lines: A Systematic Review," SPLC, San Francisco, California, 2009.

[10] Brisaboa, N.R., Cortiñas, A., Luaces, M.R. and Pol'la, M., "A reusable software architecture for geographic information systems based on software product line engineering". In *Model and Data Engineering* (pp. 320-331). Springer, Cham, 2001.

[11] Apel, S., Batory, D., Kästner, C. and Saake, G.. "*Feature-Oriented Software Product Lines: Concepts and Implementation*". Springer-Verlag Berlin Heidelberg, 2013.

[12] M.Snoeck, S.Poelmans, G.Dedene, "A Layered Software Specification Architecture" , 2000.

[13] Amina Guendouz, Djamal Bennouar, "Component-Based Specification of Software Product Line Architecture", 2014.

Weighted Distance-Based: An Enhanced String Technique for Test Case Prioritization

Muhammad Khatibsyarbini, Mohd Adham Isa and Dayang Norhayati Abang Jawawi
School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia,
81310 Johor Bahru, Johor, Malaysia
[e-mail: fkmuhammad4@gmail.com]
*Corresponding author: Muhammad Khatibsyarbini

Abstract

This paper presents an enhanced string distance metric with weighted string technique. The proposed technique applied in test case prioritization (TCP) environment. The technique aims to improve the distances priority between test cases and to reduce distance redundancy. For this work, three datasets of a similar application which is *tcas* program, but different version of language is utilized. From the datasets, using the test cases available, the differences among them were calculated using the string distance metric. In order, to improve priority and reduce redundancy, weight element was added resulted as enhanced weighted distance-based string technique. Subsequently, using the generated distance among the test cases, a simple prioritization algorithm was used to identify which set of test cases have the best combination for TCP. The results of this experiment were calculated using the most utilized evaluation in TCP which average percentage fault detection (APFD).

***Keywords*:** String Distance, String Metric, Test Case Prioritization, Software Engineering

1. Introduction

String metric or textual examination metric is a technique of finding the closeness separation of two arrangements of content [1], [2]. It is generally utilized as a part of record linkage where a quick and effective method for ascertaining general comparability between two records is required for expansive informational collections. The comparability is characterized in light of specific prerequisites rather than a solitary trait of the records. Text similarity measures play an increasingly important role in text related research and applications in tasks such as information retrieval, text classification, document clustering, topic detection, topic tracking, questions generation, question answering, essay scoring, short answer scoring, machine translation, text summarization and others [3].

There were two major based in string metric; distance-based and weighted-based string metric [4]. For each based have two type of calculation ideas where for distance-based were calculate using character or term while for weighted-based using term and token [5]. The summary of this diferrences shown in **Fig. 1** which illustrate the string metric taxonomy.

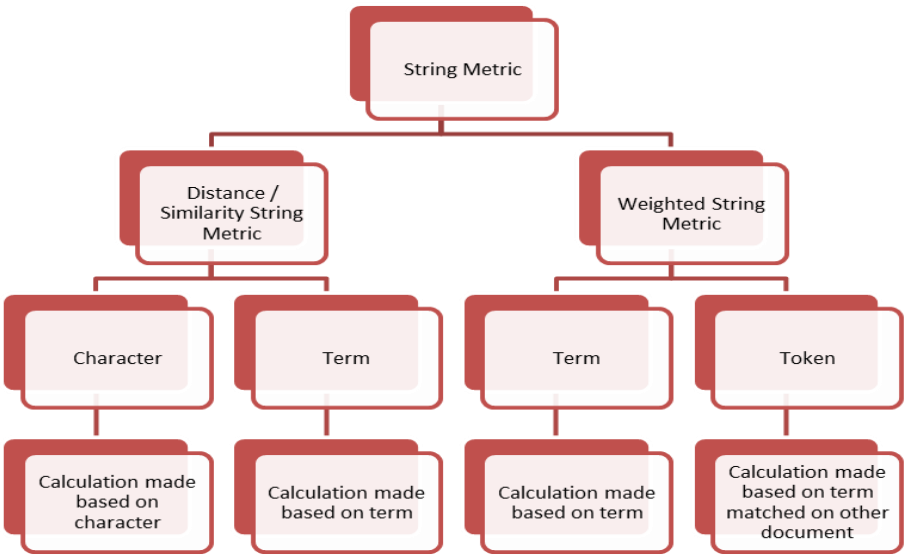


Fig. 1. Overview of String Metric Category

The strategy for each string based are different. As for distance-based, the strategy is to differentiate between two or more information. In TCP context is the test cases differences [1], [6]. While for weighted-based, the strategy is to weighted the information to give value or priority. In TCP context is the test cases importance. There were numerous works on string metric within TCP context however there were no solid work which focus on enhancement of string metric with consideration of both type of string metric [6]–[10]. This implies an enhancement for string distances metrics with other related metrics such as weigh based were worth for further analysis.

2. Overview of Basic Distance-Based String Technique

The strategy of basic distance-based string metric is to differentiate between two or test information or test cases in the context of TCP in the view of two dimensional calculations[1], [11]. The calculations were made solely to find the distance among them. **Fig. 2** illustrate the example results of distance-based calculations and plotted on two dimensional chart.

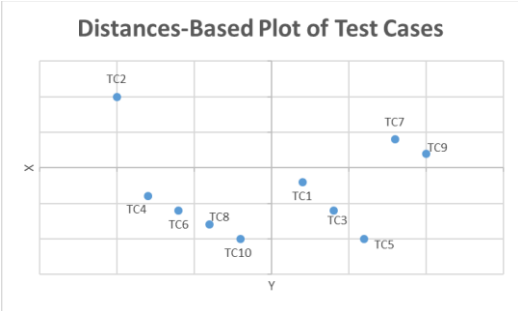


Fig. 2. Illustration of basic distance-based string technique

From **Fig. 2**, the result of distance-based string metric were plotted in two dimensional chart since the main strategy focus on the distance only. As the result, this strategy have cause some problems in TCP such as the lack of priority of test case and there will be high number of redundancy of similar distance which make it difficult to prioritize later on [12]–[15].

3. Overview Idea of Enhanced Weighted Distance-Based String

As the distance-based were made based on the differences between two points, instead of using one metric only which is distance to differentiate the test cases, why not to add other type of metric such as weight-based string metric to the point as second metric to be consider. This had given the

inspiration to enhance the string distance-based with weighted-based for text document to solve the problem regarding the priority between test cases and to reduce distance redundancy. **Fig. 3** illustrate the example results of enhanced weighted distance-based plotted chart.

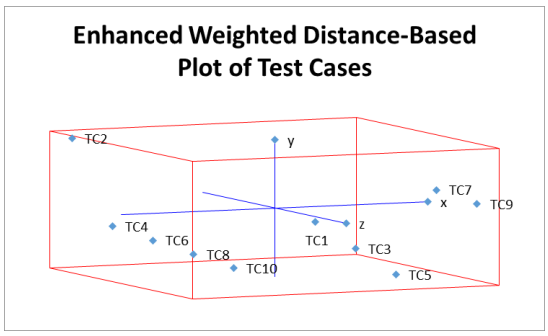


Fig. 3. Illustration of enhanced weighted distance-based string technique

From **Fig. 3**, the result of distance-based string metric were plotted in three dimensional chart as the weighted element has been added which provide extra point of pivot. This extra point of pivot may give significant improvement to counter the distance redundancy as well as the priority problem. From the three dimension view also could benefit in providing new technique enhancement in TCP context such as having classification and clustering technique before executing TCP process itself.

4. Enhancement Steps and Modification

In order to make the enhancement, two string metric were chosen for this experiment. Levensthein string metric from string distance-based and term frequency–inverse document frequency (TF-IDF) for string weighted-based. The overview idea of enhanced string distance formulation is shown in Figure 4. From the figure the enhanced string distance process start with the selection string metrics to be used. Then, the weight-based string metrics which is TF-IDF were then separated into two parts. The TF undergo changes to find the frequency differences between two document which represent the weight of the link between them. However, for IDF part does not undergo any modification, only the place of IDF formulation were made adjustment to fit in into new TF. This allows the term frequency and inverse document frequency tally with the distance values calculated.

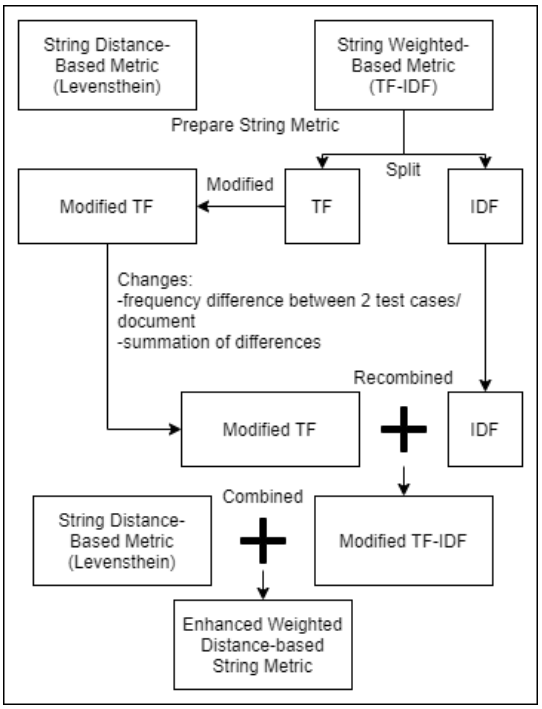


Fig. 4. The enhanced string metric

5. Experiment Setup and Design

For this experiment three datasets of a similar application which is tcas program, but different version of language is utilized. Table 1 shows a clear overview of the datasets used.

Table 1. Overviews of Datasets

Dataset Name	Programming Language	Fault Matrix
TCAS	C	Different Fault Adjacency Matrix
JTCAS	JAVA	
CSTCAS	C#	

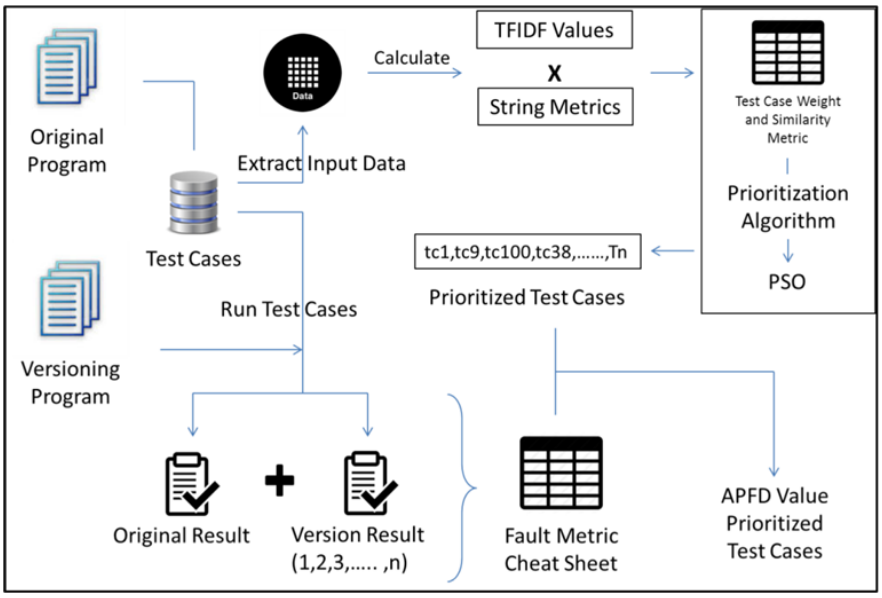


Fig. 5. Experiment framework design

From Table 4.4, the dataset used is TCAS which is originally a C program of an aircraft collision avoidance system. It takes 12 integer inputs and produces one output. The program came with one base version and 41 faulty versions with 1608 test cases. The fault matrix is produced by executing all test cases on all 41 faulty versions and compared against their base version. Even as all the datasets are for the same application, the fault matrix for each version of the application is different. The fault adjacency matrix is obtained based on calculated distances among test cases. To execute the experiment, **Fig. 5** shows the overview of the experiment framework design.

From **Fig. 5**, the experiment can be divided into three phases; Information Extraction phase, String Distance and Prioritization phase, and Evaluation phase. For information extraction phase, the phase starts with extracting all the test cases and their inputs. The extracted inputs for each test case were put into different text document to ease the calculation needed in the next phase. Then, the original and versioning programs were run through the extracted test cases to get their respective result. The results for each version were compared with original result to produce fault matrix sheet to be used in evaluation phase.

The next phase is the calculation of string distance and prioritization phase. The string distances between test cases are calculated based on the extracted inputs and populated into test case distance matrix. Within this calculation part, each test case is differentiated from each other based on the similarity distance percentage of all string metrics. The proposed enhanced string distance is also carried out here. Upon completion of the calculation, the test case similarity distance percentage matrix is then populated and prioritized using simple nearest neighbor algorithm. Until this phase, the enhanced string distances will be evaluated by calculating the APFD and coverage effectiveness (CE) values.

Finally, in the evaluation phase, the prioritized test suites are evaluated by calculating their APFD rate based on fault matrix of their program. The fault matrix was obtained in earlier stage by executing all test cases for all versioning programs and comparing the outputs against the original program. The results are compared to produce the fault matrix sheet. The CE value also calculated and converted into percentage value to ease the comparison between the results.

6. Results and Discussion

As for the results, the enhance weighted distance-based string metric show a promising result. **Table 2** show the overall results for APFD and CE from the experiment. The table compare the result for the non-enhanced distance-based string metric which is Levensthein string metric with the enhanced weighted distance-based string with TF-IDF.

Table 2. APFD and CE results for TCAS experiment

		Non Enhanced Distance-Based String Metric	Enhanced Weighted Distance-Based String Metric
APFD	tcas	92.39	**95.87**
	jtcas	93.80	**96.44**
	cstcas	93.06	**95.56**
CE	tcas	7.78	**62.38**
	jtcas	8.46	**63.49**
	cstcas	7.92	**69.82**

From Table 2, the results clearly show that the enhanced string metric perform better than the non-enhanced one. The enhanced string metric show slight improvement for APFD results in all three datasets. As for coverage effectiveness, the result surprisingly show significant improvement as with more that 50% improment for each datasets. This shows that the enhancement does solve the priority problem lies within test case apart from remove the redundancy of distances.

7. Conclusions

Based on the result it can be conclude that the idea of weighing TF-IDF is to calculate the weight of each test case based on the frequency of term used, give extra priority to the input values within the test case based on their occurrences has resulted in better APFD and CE as compared to non-enhanced formulation. As for future work, it would be quite interesting to have this enhancement applied into other distance-based string metric to see which combination would give better results.

References

[1] Y. Ledru, A. Petrenko, S. Boroday, and N. Mandran, "Prioritizing test cases with string distances," *Autom. Softw. Eng.*, vol. 19, no. 1, pp. 65–95, 2012.

[2] A. Shahbazi and J. Miller, "Black-Box String Test Case Generation through a Multi-Objective Optimization," *IEEE Trans. Softw. Eng.*, vol. 42, no. 4, pp. 361–378, 2016.

[3] M. Khatibsyarbini, M. A. Isa, D. N. A. Jawawi, H. N. A. Hamed, and M. D. Mohamed Suffian, "Test Case Prioritization Using Firefly Algorithm for Software Testing," *IEEE Access*, vol. 7, pp. 132360–132373, 2019.

[4] M. Khatibsyarbini, M. A. Isa, and D. N. A. Jawawi, "A hybrid weight-based and string distances using particle swarm optimization for prioritizing test cases," *J. Theor. Appl. Inf. Technol.*, vol. 95, no. 12, 2017.

[5] W. Gomaa and A. Fahmy, "A survey of text similarity approaches," *Int. J. Comput.*, 2013.

[6] M. Khatibsyarbini *et al.*, "Test case prioritization approaches in regression testing: A systematic literature review," *Inf. Softw. Technol.*, vol. 93, pp. 74–93, Sep. 2018.

[7] B. Jiang and W. K. Chan, "Input-based Adaptive Randomized Test Case Prioritization," *J. Syst. Softw.*, vol. 105, no. C, pp. 91–106, 2015.

[8] C. Fang, Z. Chen, K. Wu, and Z. Zhao, "Similarity-based test case prioritization using ordered sequences of program entities," *Softw. Qual. J.*, vol. 22, no. 2, pp. 335–361, 2014.

[9] M. L. Mohd-Shafie, W. M. N. Wan-Kadir, M. Khatibsyarbini, and M. A. Isa, "Model-based test case prioritization using selective and even-spread count-based methods with scrutinized ordering criterion," *PLoS One*, vol. 15, no. 2, p. e0229312, Feb. 2020.

[10] M. Abdur, M. Abu, and M. Saeed, "Prioritizing Dissimilar Test Cases in Regression Testing using Historical Failure

Data," *Int. J. Comput. Appl.*, vol. 180, no. 14, pp. 1–8, 2018.

[11] D. Gusfield, *Algorithms on strings, trees, and sequences : computer science and computational biology*. Cambridge University Press, 1997.

[12] B. Jiang, Z. Zhang, W. K. Chan, and T. H. T. Tse, *Adaptive random test case prioritization*. 2009, pp. 233–244.

[13] 2 1R.Kanimozhi and Prof.J.RBalakrishnan, "Cosine Similarity Based Clustering For Software Testing Using Prioritization\n," *IOSR J. Comput. Eng.*, vol. 16, no. 1, pp. 75–80, 2014.

[14] J. Zhou and D. Hao, "Impact of Static and Dynamic Coverage on Test-Case Prioritization: An Empirical Study," in *2017 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, 2017, pp. 392–394.

[15] S. Yoo and M. Harman, "Regression testing minimization, selection and prioritization: A survey," *Softw. Test. Verif. Reliab.*, vol. 22, no. 2, pp. 1–7, 2012.

Analysis of Level of Service on Sewage Treatment Facility's Efficiency Improvement and Electricity Reduction

Soyeong Park, Sungmin Ko and Jungwoo Shin*
Department of Industrial Management Engineering, Kyunghee University
Suwon, South Korea
[e-mail: sysy@khu.ac.kr , sungmine03@khu.ac.kr, shinjung11@khu.ac.kr]
*Corresponding author: Jungwoo Shin

Abstract

In order to reduce GHG emissions for a sustainable environment, we have set GHG reduction targets around the world, and movements to realize them are continuing. The waste sector is regarded as one of the main causes of greenhouse gas emissions, and the improvement of energy efficiency of sewage treatment facilities is being given as a task. In this study, the service level for sewage treatment efficiency is derived through consumer preference analysis for sewage treatment facilities. Therefore, we will discuss the expected electricity reduction effect through this, and present the direction of asset management through DB for sewage treatment facilities.

***Keywords*:** Energy reduction, Energy efficiency improvement, Asset management, Sewage treatment facility

1. Introduction

Water resources are an important part of public health and the environment. It helps to improve economic development and quality of life[1]. The United Nations projects the world's population to reach 8.5 billion in 2030, 9.7 billion in 2050 and 10.9 billion in 2100[2]. As the population grows, water consumption will gradually increase, and as a result, the enormous amount of sewage becomes an important issue. Sewage treatment is an expensive process. Sewage management requires enormous cost, and nutrients released during treatment cause eutrophication of surrounding aquatic ecosystems[3]. Therefore, it is necessary to improve the efficiency of sewage treatment and manage costs well in order to preserve the ecosystem and reduce urban management costs.

There are studies on sewage treatment technology for cost-effective industrial wastewater treatment[4]. In addition, technical studies for low-cost sewage treatment systems are being conducted[5, 6]. There is also a study on the operation and management of sewage treatment facilities to reduce the cost of sewage treatment[7]. In addition, research on the overall water infrastructure management method and the development of a model for water resource asset management were also conducted[8, 9].

However, there are not many studies on the effect of reducing electricity consumption by improving sewage treatment efficiency.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2017H1D8A2031138).

Therefore, this study analyzes the power reduction effect according to the sewage treatment efficiency through consumers' willingness to pay for sewage services. It also proposes a direction for asset management based on consumer preference.

Chapter 1 will discuss the background and purpose of this paper, and Chapter 2 will discuss the methodology used in this paper. Chapter 3 discusses the collection and organization of data, the results of analyzing the data in Chapter 4, and the conclusion in Chapter 5.

2. Methodology

CVM and CE are representative methods of consumer preference analysis. CVM has the advantage of being able to measure the value of non-market goods through scenarios, but since CE can grasp the value of attributes through alternative cards, this study investigated consumer preferences through CE.

Data was collected on consumer preferences for alternatives. MNL was used as an analysis method, but assuming the distribution of the error term as iid Type Ⅰ extreme value distribution, the substitution relationship between alternatives did not change despite the introduction of a new alternative. Based on this, it is possible to understand consumer preferences, so this study intends to analyze it through MNL[10].

The probability P that consumer n chooses the i alternative out of the j alternative is given by Equation (1)[10]:

$$P_{ni} = \int \left(\frac{\exp(\beta_n' x_{in})}{\sum_j \exp(\beta_n' x_{jn})} \right) f(\beta) d\beta \quad . \qquad \textbf{(1)}$$

Marginal WTP (MWTP) and relative importance (RI) are calculated through coefficients of estimated variables. MWTP refers to the amount paid by respondents as the attribute level changes. RI refers to the degree of impact each attribute has on the choice of an alternative by comparing the magnitude of the utility given to the attribute. Equations (2) and (3) mean this[11]:

$$\mathrm{MWTP}_{x_{kj}} = -\frac{\beta_j}{\beta_{k,price}} \qquad (2)$$

$$\mathrm{RI}_k = \frac{part\ \ worth_k}{\sum_k part\ \ worth\_k} \times 100 \quad . \qquad (3)$$

3. Experiment

3.1 Data

This study used SP (Stated preference) data to analyze consumer preference for sewage treatment facilities. The survey was conducted in September 2020, and 224 people were surveyed online in Korea. The survey was conducted by Gallup Korea, an online survey organization. **Table 1** below shows the demographic distribution of the respondents who were subject to this survey.

Table 1. Demographic properties of sample

Category	Characteristic	Respodents (n)	Percentage (%)
Total		224	100
Gender	Male	116	51.8
	Female	108	48.2
Age	20-29 years	44	19.6
	30-39 years	42	18.8
	40-49 years	47	21.0
	50-59 years	49	21.9
	60-69 years	42	18.8
Average monthly income per household	Under KRW 2 million	31	13.8
	KRW 2-3 million	29	12.9
	KRW 3-4 million	34	15.2
	KRW 4-5 million	32	14.3
	Over KRW 5 million	98	43.8

This questionnaire consisted of total three parts, and the first part consisted of questions about respondents' overall prior knowledge about sewage treatment facilities and satisfaction with current services. In the second part, the preferred type of sewage service was selected through the configured alternative card, and the third part

was composed of respondents' demographic questions.

The attributes of the alternative card are: First, the conditions for hydrophilic and leisure activities, which means the range of water-friendly activities due to the sewer service, second, the level of odor caused by the sewer service, and third, due to insufficient sewage capacity during rainfall Fourth, sewage treatment efficiency, which is the level at which sewage is transported to the sewage treatment plant, and fifth, total time required for handling civil complaints for sewage, and finally, monthly average sewerage rates are presented. **Table 2** shows the attributes and levels determined by this study.

Table. 2. Attributes and attribute levels

Attributes	Attribute levels
Conditions for water-friendly leisure activities	- Good - Normal - Caution
Odor level	- 1 degree of odor intensity - 2 degree of odor intensity - 2.5 degree of odor intensity
Inland flooding	- 1 case/year - 3 case/yea - 5 case/yea
Efficiency of sewage treatment	- over 90% - 70%-90% - 50%-70% - under 50%
Complaint processing time	- under 2hours - 2hours-5hours - 5hours-30hours - over 30hours
Sewerage Fee	- KRW10,000 - KRW 20,000 - KRW 30,000 - KRW 40,000

Considering all the number of attribute cases, the total was 1,728 (3*3*3*4*4*4=1,728). Through the Statistical Package for the Social Sciences (SPSS) through orthogonal fractional factorial design, 32 alternative cards were presented to respondents.

3.2 Result

This study was analyzed through a mutinomial logit model. The utility equation for estimation is as follows.

$$U_{nj} = \beta_1 X_{Activity} + \beta_2 X_{Odor} + \beta_3 X_{Flooding}$$

$$+\beta_4 X_{Efficiency} + \beta_5 X_{Time} + \beta_6 X_{Fee} + \varepsilon_{nj}$$

Consumers tend to prefer the better the water-friendly leisure activity conditions, the lower the odor level, and the less domestic flooding occurs, the higher the sewage treatment efficiency, and the shorter the civil petition processing time is. In addition, it was found that the results suitable for the reality were obtained by preferring to have a low sewage fee.

The relative importance was highest in the order of sewage rate, odor level, and hydrophilic leisure activities. In addition, he was willing to pay about KRW19,000 to prevent the odor level from rising to the next level, and about KRW15,000 for a better environment compared to the case where the conditions of the pro-benefit activities are normal.

In order to improve the efficiency of sewage treatment to the next level, it was willing to pay an additional monthly sewage fee of about 3,000 won. The number of households in Korea is approximately 21 million in 2019, and if the willingness to pay to improve the efficiency of sewage treatment is expanded, it is equivalent to KRW 69.4 million.

In addition, it can be expected to reduce power by improving the efficiency of sewage treatment. The benefits from energy reduction can be utilized for infrastructure improvement and sewage treatment technology development.

4. Conclusions

This study analyzed consumer preference and power savings for the efficiency of sewage treatment facilities through CE. In sewage treatment facilities, sewage rates, odor levels, hydrophilic leisure activities, and sewage treatment efficiency were important attributes.

The willing amount to pay for sewage treatment efficiency is about KRW 69.4 million per year, and power reduction effects can be expected due to sewage treatment efficiency. As the financial support is definitive for the effect of the reduction in power use due to the increase in efficiency, this will lead to a virtuous cycle of reducing GHG through technology development and infrastructure construction and improvement.

In addition, it is possible to manage assets by establishing a sewage treatment facility DB. The quality of service can be improved by predicting problems arising from aging faciliities and supplementing the items that consumers prioritize among service items when a problem occurs.

In the future, future studies on the integration of people's willingness to pay, capacity for service improvement, and electricity bills according to the improvement of sewage treatment efficiency should be conducted, and an in-depth study on why payment rejectors refused.

References

[1] K. Stamatelatou, K. Tsagarakis, “Sewage treatment plants: Economic evaluation of innovative technologies for energy efficiency,” *IWA Publishing*, Vol 1, 2015.

[2] United Nations, Department of Economic and Social Affairs, Population Division "World Population Prospects 2019: Highlights,” 2019.

[3] P. Bolund, S. Hunhammar, “Ecosystem services in urban areas,” *Ecological economics*, Vol 29. No. 2, pp. 293-301, 1999.

[4] R. Abma, W. Driessen, R. Haarhuis, et al., “Upgrading of sewage treatment plant by sustainable and cost-effective separate treatment of industrial wastewater,” *Water science and technology*, Vo. *61, No.* 7, pp. 1715-1722, 2010.

[5] M. Tandukar, S. Uemura, I. Machdar, et al., "A low-cost municipal sewage treatment system with a combination of UASB and the “fourth-generation” downflow hanging sponge reactors." *Water science and technology,* Vol.52, No. 1-2, pp. 323-329, 2005.

[6] I. Machdar, H. Harada, A. Ohashi, et al. "A novel and cost-effective sewage treatment system consisting of UASB pre-treatment and aerobic post-treatment units for developing countries." *Water Science and technology* Vo. 36, No. 12, pp.189-197, 1997

[7] H. Kang, "Challenges for water infrastructure asset management in South Korea." *Water Policy* Vol. 21, No. 5, pp.934-944, 2019.

[8] H. Park, M. An, N. Cho, et al., “Improvement for asset management of sewage treatment facilities,” *Journal of the Korean Society of Civil Engineers*, Vol. 31, No. 2D, pp. 285-293, 2011.

[9] A. Heather, J. Bridgeman, "Water industry asset management: a proposed service- performance model for investment." *Water and Environment Journal,* Vol.21, No. 2, pp. 127-132, 2007.

[10] T. Kenneth E, “Discrete choice methods with simulation,” *Cambridge university press*, 2009.

[11] C. Berglund, "The assessment of households' recycling costs: The role of personal motives," *Ecological Economics*, Vol 56, No. 4, pp. 560-569, 2006.

Impact Analysis of Change Event on Execution Time of IT Service Process

Hoonseok Park[2] and Jae-Yoon Jung[1,2]
[1]Department of Industrial and Management Systems Engineering, Kyung Hee University, South Korea
[2]Department of Big Data Analytics, Kyung Hee University, South Korea
[e-mail: hoonseok@khu.ac.kr, jyjung@khu.ac.kr]
*Corresponding author: Jae-Yoon Jung

Abstract

IT services require continuous improvement and management of various components such as products, services, and processes to meet the changing needs of their customers. This study proposes a method to analyze the effect of change events on execution time of ITIL-based service process. First, the start time and end time of change events were compared with those of the IT service process, and then the types of cases were categorized. Next, process mining and statistical techniques were used to analyze the effect of change events on service process execution time and extract the process map of service processes corresponding to each type. The ITIL-based process data provided by BPIC 2014 were used to evaluate the proposed method. As a result, it was confirmed that the change event of the configuration items did not affect the business performance of the service desk team, but the business performance of the IT operation team.

***Keywords*:** Information technology infrastructure library (ITIL), process mining, IT service process, change management

1. Introduction

The Information Technology (IT) plays a crucial role of supporting business value creation. IT service management can be based on the Information Technology Infrastructure Library (ITIL) framework. ITIL is a standard framework that presents best practices for establishing systematic design, operation, and improvement of IT services [1].

ITIL proposes a process structure to support high quality of IT services. Changes of configuration items cause changes in the function of the service process, and may affect the negative performance of the IT service process.

In this study, we proposes a method of analyzing the effect of change events on the execution time in IT service processes based on the ITIL framework. In particular, process mining [2] and statistical techniques were used to analyze the effect of change events on service process execution time and extract the process map of service processes corresponding to each type.

The remainder of this paper is structured as follows. In Section 2, we propose the proposed method based on the relationship between change event and service process. In Section 3, the proposed method is evaluated using the ITIL process data provided by BPIC 2014. Finally, Section 4 concludes this paper with future research.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2017H1D8A2031138, 2019R1F1A1064125).

2. Method

In this study, a change event is assumed to be an event of IT service components that causes the change in the state or attribute of one or more certain subjects in the service process. For example, in the case of an IT service company, items such as servers and PCs can be the important subject of IT service, and updates and improvements to these items may correspond to change events.

Change events can occur in various patterns depending on the execution time that the process is performed. **Fig. 1** shows the types of patterns that can occur between change events and process cases. Each case related to the process is processed sequentially through several events over time. For example, when the inquiry of a customer can initiate a single process case, the case is performed through events such as email confirmation, inquiry processing, and email reply. The treatment of change events is related to the activity execution in the service process, and can be linked to each other in a data structure.

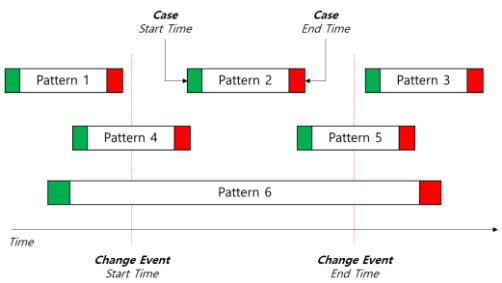


Fig. 1. Patterns between event change and service process

Patterns between change events and process cases are categorized into six types. The patterns reveal six possible relationships that can occur between the start and end times of the change event, and those of the corresponding process case. It can be seen that process cases can be executed before and after a change event occurs or while a change event is being executed.

It is necessary to analyze the six pattern types between the change events and processes. The method analyzes process cases belonging to each pattern from a statistical point of view and a process workflow point of view.

3. Experiments

The ITIL-based process event data of the business process intelligence challenge (BPIC) 2014 [3], provided by the Rabobank ICT Group, was used for this experiment. The experimental data provides the Change events, the Interaction history of the service desk team, the Incident history of the IT operation team, and the Incident Activity history that recorded the detailed events related to the Incident.

3.1 Data Preprocessing

In the *Interaction* and *Incident* data, there were cases that were not completely closed. To prevent the edge effect caused by these cases, only cases in which the execution status is 'Closed' were selected. In the *Change* data where the change events is recorded, the 'Actual Start' column is selected for the start time and the 'Actual End' column is selected for the end time. If the columns have missing values, the values of the 'Planned Start' and 'Planned End' columns were used instead.

After that, the *Change* data and the *Interaction* and *Incident* data were joined by using the 'CI Name' column, which is the unique number of the configuration item. After going through all the pre-processing, we were able to prepare two combined data for analysis, *Change-Interaction* and *Change-Incident*.

Table 1. Sample data of Change-Interaction cases

CI Name	Change ID	Start Time	End Time	Inter-action ID	Open Time	Close Time	Handle Time (sec.)
HMD 000002	C000 00003	2013 -12-18 14:00	2013 -12-18 16:15	SD00 75751	2014 -01-03 11:45	2014 -01-08 15:24	426
HMD 000002	C000 00003	2013 -12-18 14:00	2013 -12-18 16:15	SD01 21418	2014 -02-26 11:28	2014 -03-31 16:00	0
ASW 000010	C000 00007	2013 -06-21 09:00	2013 -09-02 18:00	SD00 00657	2013 -09-11 09:48	2013 -10-04 09:20	151
ASW 000010	C000 00007	2013 -06-21 09:00	2013 -09-02 18:00	SD00 01046	2013 -10-07 14:19	2013 -10-04 09:58	165
ASW 000010	C000 00007	2013 -06-21 09:00	2013 -09-02 18:00	SD00 07223	2013 -10-10 14:12	2013 -10-07 16:46	265

Table 1 shows the structure and sample data of the Change-Interaction dataset. It includes the start time and end time of change events and service processes, along with the 'Handle Time'

column that records the execution time to be used for effect analysis of change events.

3.2 Impact Analysis of Change Events

The patterns were used to investigate the effect of change events on service processes. Multiple regression analysis was applied to two datasets to test whether there was a statistically significant difference in process execution time according to the pattern type. The results of multiple regression analysis showed that change events does not affect statistically significantly the execution times of the Change-Interaction cases of the service desk team at the level of 0.05. Also, it was the same to the Change-Interaction cases of the IT operation team.

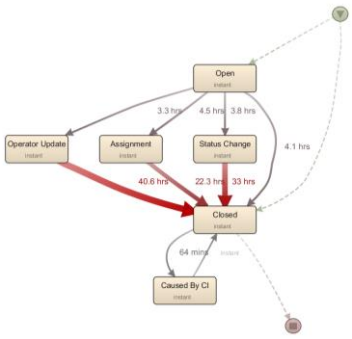


Fig. 2. The execution time of the change-interaction cases of pattern 1

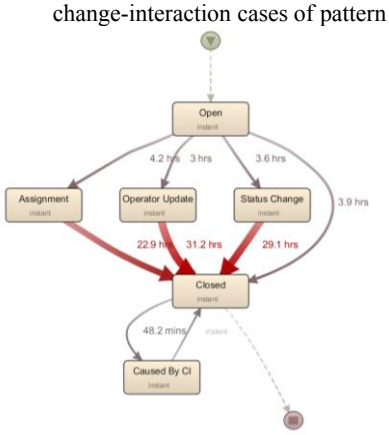


Fig. 3. The execution time of the change-interaction cases of pattern 3

Fig.s 2 and **3** show the process maps of patterns 1 and 3 with their execution time, respectively. They were discovered by using Disco, which is a process mining tool of Fluxicon [4]. In summary, pattern 1 represents a case of a service process ended before a change event started, and pattern 3 represents a service process case started after a change event was over. For the comparison of cases by pattern, the activity filter was set to 10% and the path filter was set to 0%. Among the events that had executed between 'Open' and 'Closed', the 'Operator Update' showed the most different execution time between the two processes by 9.4 hours, and the time decreased in pattern 3. In addition, it was confirmed that the time of 'Caused By CI' was an average of 64 minutes before the change event, but decreased to an average of 48.2 minutes after the change event was performed. Therefore, it can be said that the change event for the configuration item affected the 'Operator Update' and the 'Caused By CI' from the process point of view.

4. Conclusions and Future Work

In this study, a method was proposed to analyze the effect of change events in IT service processes based on the ITIL framework. To do that, regression analysis was used in the statistical point of view, and time analysis was used in the process point of view. The proposed method was illustrated with the IT service process data provided by BPIC 2014.

To define the relationship between change events and service processes, this study utilized the start and end times of events and cases. However, there may be potential factors such as time as well as resources and change events previously performed. As future work, it is necessary to define and analysis different patterns considering the other factors.

References

[1] A., Valerie, ITIL®: The Basics, 2010.
[2] W. van der Aalst, "Process mining – Data science in action," Springer, Berlin, Heidelberg, 2016.
[3] IEEE Task Force on Process Mining, "Business Process Intelligence Challenge (BPIC)," The 10th International Workshop on Business Process Intelligence, 2014, http://www.win.tue.nl/bpi/2014/challenge
[4] Fluxicon, Automated Process Discovery Software for Professionals - Fluxicon Disco, 2020. http://fluxicon.com/disco/

Forecasting of Energy Demand in the Railway Sector Using Machine Learning

Myoungjin Oh[1], Hanseul Jo[2] and Jungwoo Shin[1*]
[1]Department of Industrial and Management Systems Engineering, Kyung Hee University, South Korea
[2]Department of Applied Bigdata systems Engineering, Kyung Hee University, South Korea
[e-mail: bristo94@khu.ac.kr, johansle@khu.ac.kr, shinjung11@khu.ac.kr]
*Corresponding author: Jungwoo Shin

Abstract

This study was conducted to forecast Korea's future energy demand in the railway sector. KNN regression was used as the analysis method, and ASIF was assumed for the overall model. In order to forecast energy demand in the railway sector, the KNN regression and ARIMA results were compared, and it was found that the forecasting power of KNN regression was higher than ARIMA. Based on the estimation results, energy consumption is predicted by 2025, and the forecasting results shows that overall energy consumption increases in each year. However, while energy consumption of diesel has increased, electric consumption has decreased.

***Keywords*:** Railway Sector; Energy Demand; Machine Learning; KNN regression;

1. Introduction

According to the IEA[1], 23% of global greenhouse gas emissions come from the transport sector. As Korea has signed the Kyoto Protocol, it must reduce its carbon emissions by 2030, so efforts to reduce emissions must continue in the transport sector. To this end, it is important to convert the demand from the transport sector to the rail sector, one of intermodal transport, which is an eco-friendly means of transport[2, 3]. This study attempts to predict and interpret the future energy consumption of the railway sector in Korea. As a methodology for this, we use KNN regression, one of machine learning techniques.

2. Method

The energy use forecast methodology of the railway sector in this study assumes ASIF (Activity → Structure → Intensity → Fuel) as the basic form. In the case of activity, it is divided into passengers and freight. In the case of railway sector, diesel and electric power are mainly used, so they are classified through the structure. In other words, the share of diesel passengers is calculated as the diesel passenger estimate/(diesel passenger estimate + electric power passenger estimate). Intensity refers to fuel economy, and Fuel refers to energy consumption. All parts are estimated by KNN.

Machine learning techniques, which have developed greatly along with the recent development of computer computing capabilities, have expanded their scope to time series analysis. One of the frequently applied techniques is KNN regression[4, 5, 6]. KNN began with the application of Fix & Hodge[7] as a nonparametric classification method, and is a method applied in the form of a regression equation rather than a classification method.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2017H1D8A2031138).

After converting a feature of the input data into a distance, the data value is predicted through the weighted average of the data having the closest distance. In addition, after searching for K pieces of surrounding data based on the distance, a predicted value is determined by weighting the surrounding data. KNN is useful in predicting the future of the current data through the predicted values of the past data that are most similar to the current data, has high computational efficiency, and has the advantage of being robust even for noisy training data (Kuck & Freitag, 2020).

In this study, K and the unit of time are determined by verifying the data through 3-year data. In the case of activity and structure, monthly data is used, so K (1 to 35) and time unit (3 to 36) are determined by comparing actual data and predicted data from January 2016 to December 2018, the latest data. In the case of intensity, on the other hand, since annual data is used, K (1 to 35) and time unit (3 to 12) are determined by comparing actual data and predicted data from 2016 to 2018, the most recent data. All distances were Euclidean distances, and LOOCV (Leave-One-Out Cross Validation) was used as the verification method. All estimates were analyzed with monthly GDP(by denton[8]) through the OECD industrial production index), population and oil prices interpolated. At this time, the verification power is judged through RMSE.

3. Experiments

3.1 Data

In this study, for activity, the Converted car-km (passenger car, unit: km) and Freight Traffic Ton-km (unit: ton-km) of Statistical yearbook of railroad were used, and the available data were from January 1987 to December 2018. In the case of structure, Diesel Locomotive, Diesel Rail Car were summed as diesel, and Electric Locomotive, High speed railway, SMESRS, Electric Motor Car were summed as electric, and available data are the same as activity. Since intensity data is absent, km/toe and ton-km/toe data were generated using railway statistics yearbook and yearbook of energy statistics, respectively, and available data are from 1987 to 2018.

In addition, quarterly GDP can be obtained through the Bank of Korea, the OECD industrial production index, population and oil prices through the Office of Statistics.

3.2 Result

Table 1 shows the estimation capability of KNN regression, and the ARIMA estimation results are reported together for comparison. ARIMA was analyzed as the same explanatory variable as KNN regression. As a result of the analysis, it can be seen that the error of KNN is small in all sectors. In particular, the difference between ARIMA and error was remarkable in the structure section.

Table 1. Error comparison between ARIMA & KNN.

Type	ARIMA		KNN	
	MAPE	RMSE	MAPE	RMSE
Passenger Car(km)	14,678,569	16,213,821	11,032,332	12,685,742
Freight Traffic (ton-km)	289,527,778	299,278,995	78,332,690	98,763,877
Diesel Passenger Car(km)	5,128,710	7,079,438	564,985	777,874
Electirc Passenger Car(km)	16,899,514	19,669,000	6,972,210	8,226,052
Diesel Freight Car(km)	9,044,683	9,204,422	1,870,106	2,499,894
Electric Freight Car(km)	2,847,509	3,472,717	1,452,932	1,929,674

Fig. 1 shows the energy consumption of diesel and electricity. In the case of diesel, it shows a slight decrease until 2025, and in the case of electricity, it appears to be steadily rising. This is because the intencity is not change noticeable until 2025, but the activity is steadily risen.

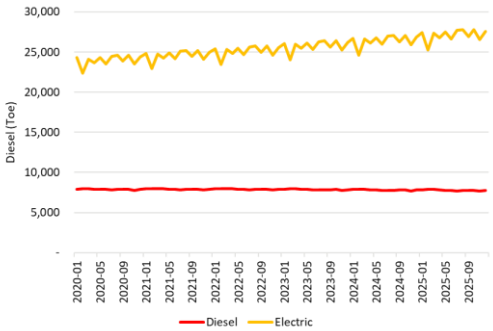


Fig. 1. Energy consumption of diesel and electric

Table 2 shows the results shown in Figure 1 by year. Energy consumption has been shown to rise steadily through 2025. This can also be interpreted as the result shown because not only the rate of increase in the energy consumption of electric power is faster than that of diesel, but also the energy share of electric power is higher than that of diesel.

Table 2. Total energy consumption per year

Year	Total energy consumption (unit: TOE)
2020	1,875,152
2021	1,894,981
2022	1,936,850
2023	1,974,282
2024	2,015,339
2025	2,058,142

4. Conclusions and Future Work

This study utilized KNN regression assuming the ASIF type to predict the energy demand of the railway sector. As a result of the analysis, it was found that energy consumption increased with time. This is because the demand for railroads has increased, while fuel efficiency has not been improved. If fuel economy improves in the railway sector in the future, energy consumption could also decrease. In addition, the increase in demand in the railroad sector can be interpreted as absorbing the demand from other transport sectors, but for a strict interpretation, additional research is needed that considers other sectors in combination.

The limitations of this study are as follows. First, the data from the railway statistics yearbook and the yearbook of energy statistics were combined to create fuel economy data. This is not accurate fuel economy information, and it has a disadvantage that it is not suitable for monthly data estimation as annual data. The second is that it has not been studied with other sectors, so it is not possible to distinguish between an increase in demand and a transfer. To compensate for this, it is necessary to use several sectors together, such as Computable general equilibrium.

References

[1] International Energy Agency(IEA), "Tracking Clean Energy Progress 2017," 2017.

[2] Black, I., Seaton, R., Ricci, A., Enei, R., Vaghi, C., Schmid, S., and Buhler, G., "RECORDIT (Real cost reduction of door-to-door intermodal transport) Final report: Actions to promote intermodal transport," *Cranfield: Cranfield University*, 2003.

[3] Kirschstein, T., and Meisel, F., "GHG-emission models for assessing the eco-friendliness of road and rail freight transports," *Transportation Research Part B: Methodological*, 73, 13-33, 2015.

[4] Ghiassi, M., Fa'al, F., and Abrishamchi, A., "Large metropolitan water demand forecasting using DAN2, FTDNN, and KNN models: A case study of the city of Tehran, Iran," *Urban Water Journal*, 14(6), 655-659, 2017.

[5] Gonzalez-Romera, E., Jaramillo-Moran, M. A., and Carmona-Fernandez, D., "Monthly electric energy demand forecasting based on trend extraction," *IEEE Transactions on power systems*, 21(4), 1946-1953, 2006.

[6] Valgaev, O., Kupzog, F., and Schmeck, H., "Low-voltage power demand forecasting using K-nearest neighbors approach," *In 2016 IEEE Innovative Smart Grid Technologies-Asia (ISGT-Asia)*, 1019-1024, IEEE, 2016.

[7] Fix, E., and Hodges, J. L., "Discriminatory analysis: nonparametric discrimination, consistency properties," *USAF School of Aviation Medicine*, 1951.

[8] Denton, F. T., "Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization," *Journal of the American Statistical Association*, 66: 99-102., 1971.

Partition-based Big Data Analysis and Visualization

Jun-Ki Hong
Department of Computer Engineering, Pai Chai University
Daejeon, South Korea
[e-mail: jkhong@pcu.ac.kr]
*Corresponding author: Jun-Ki Hong

Abstract

Nowadays, big data analytics has become very important in business and research fields and obtaining meaningful information from big data is very important. However, it is difficult to obtain meaningful information by analyzing tremendous amounts of accumulated big data. Therefore, in this paper, a partition-based big data analytic (PBDA) algorithm is proposed that calculates the representative values of partitions to obtain meaningful results from big data. The proposed PBDA algorithm can obtain meaningful analytic results from big data by using an adjustable partition size with its representative values. To verify the proposed PBDA algorithm, the sales characteristics of products are analyzed by using big data from the clothing company 'A'. By using the proposed PBDA algorithm, meaningful analysis results are obtained by analyzing the sales volume of products according to temperature change.

***Keywords*:** partition-based big data analytic, big data analysis, big data visualization

1. Introduction

A Recently, big data analytics (BDA) has become a methodology to obtain meaningful information from big data. Further, with the rapid growth of today's smart device usages, many companies collect a tremendous amount of data. Since unanalyzed big data has scalability scope and complexity, it is very important to extract and analyze meaningful information from big data. Consequently, it is becoming increasingly important to derive meaningful information from BDA to set the strategy of the companies. Various methods for big data analysis have been proposed, one of BDA techniques are partitioning big data to obtain meaningful information. Analyzing big data using partitions has the advantage of estimating the statistical characteristics of the data without calculating the entire data set [1-3]. Recently, a BDA using random sample partitions has been proposed which can estimate the characteristics of the data partitions without calculating the entire data set [4]. Furthermore, many research has been done in various fields related to the fashion industry since the BDA become very important in the fashion industry [5, 6].

The proposed PBDA algorithm divides big data into partitions to calculate the representative values. Then PBDA merges the representative values to visualize the correlations of multiple variables in the 2D graph. To verify the performance of the proposed partition-based big data analytic (PBDA) algorithm, the actual big data from clothing company 'A' is analyzed and visualized in the 2D graph. The following paragraphs explain why the big data of clothing company 'A' is analyzed.

In this study, the sales volumes of fashion goods are analyzed to obtain meaningful results from the time series big data by proposed PBDA

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2019R1G1A1100225).

2. Methodology

2.1 Data Storage

As can be seen in **Fig. 1**, when a consumer purchases a product using the Internet, the information of the date of purchase with temperature, product category and sales volume are stored in the database (DB) server since Jan. 1st, 2015.

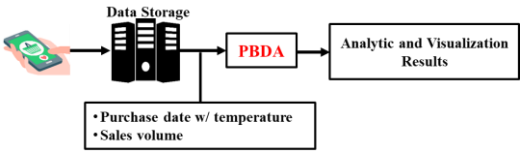


Fig. 1. Flowchart of proposed PBDA

Note that the temperature data was collected from the Korea Meteorological Administration (KMA).

2.2 PBDA

To verify the proposed PBDA, temperature with their corresponding sales volume is divided into K groups with L elements. Then, the average values representing each K-th group is obtained to analyze the correlation between temperature and sales volume.

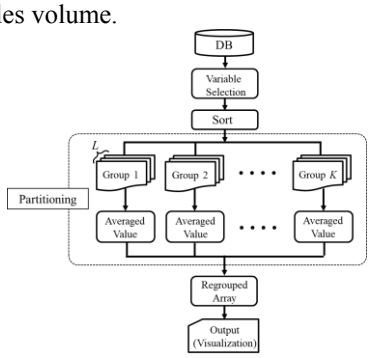


Fig. 2. Flowchart of proposed PBDA

As can be seen in **Fig. 2**, first, select and extract data for data analysis from DB. Then, sort the temperature with its corresponding sales volume in ascending order to analyze the sales volume of the product as the temperature rises. After the data is sorted in ascending order, the data are divided into K groups with L elements, partition size, and the average value of each group is calculated. Then, the calculated average values of the groups are then merged into an array to visualize the analysis results.

3. Analytic and Visualization Results

In this section, sales volume according to temperature is analyzed by the proposed PBDA for winter outers and short sleeve shirts. To analyze the sales volumes of each clothing category, the change in sales volume was analyzed by applying PBDA with non-grouping, L = 1, and grouping with L = 10 and L = 30 to verify the result changes according to the L changes.

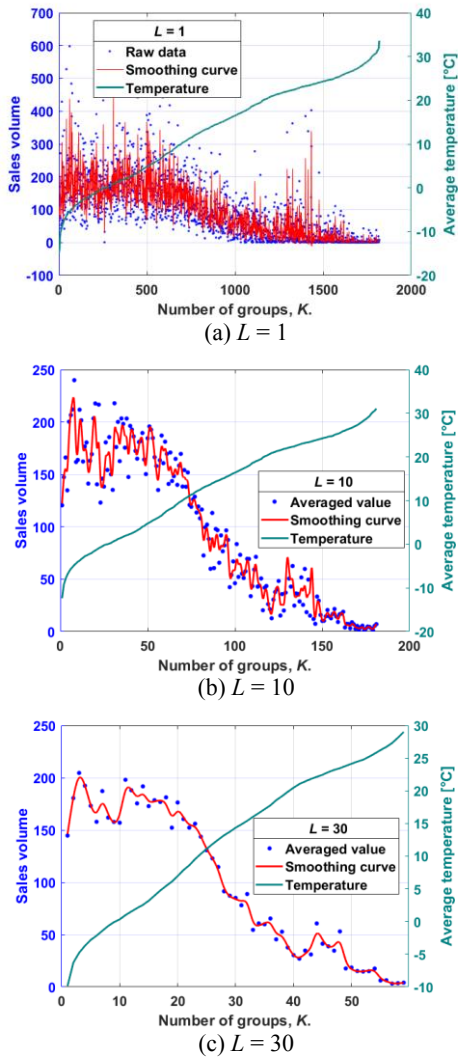
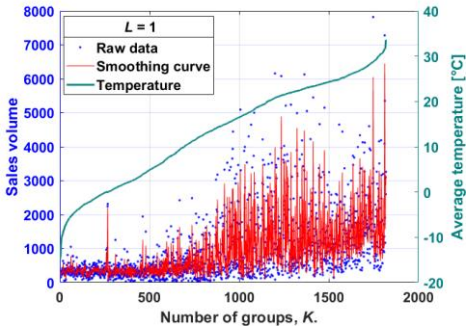


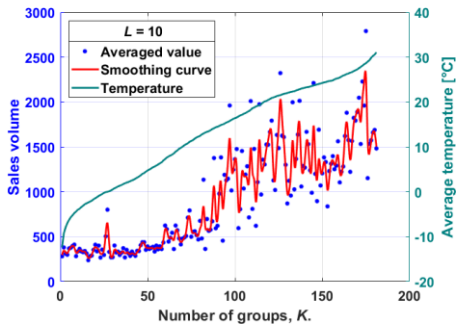
Fig. 3. Changes in sales volumes of winter outers due to average temperature change: (a) L = 1, (b) L = 10 and (c) L = 30

As can be seen in **Fig. 3**, the sales volume of the winter outers decreases as the average temperature increases. However, the analytic results shown in Figure 3 are predictable because

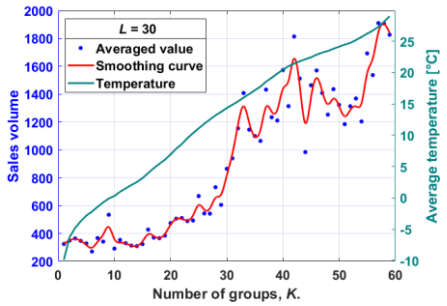
people buy winter outers as the weather gets colder. These analytic results confirm that the proposed PBDA operates correctly with the correct procedure.



(a) $L = 1$



(b) $L = 10$



(b) $L = 30$

Fig. 4. Changes in sales volumes of short sleeve shirts due to average temperature change: (a) $L = 1$, (b) $L = 10$ and (c) $L = 30$

According to the analysis results in **Fig. 4**, the sales volume of winter outers increases as the temperature decreases contrary to the analysis results for short sleeve shirts shown in Figure 3.

As a result, it is difficult to identify the change in product sales due to temperature changes only through the data sorted in ascending order with $L = 1$, but the change in sales volume due to temperature changes can be easily identified by proposed PBDA with larger L values, $L = 10$ and $L = 30$.

4. Conclusions

In this paper, PBDA is proposed to analyze and visualize the correlation between variables of the time series big data. The advantage of the proposed PBDA is that it can be implemented in any equation-computable program languages without using any specific big data analytics tools. To verify the proposed PBDA algorithm, the correlation between temperature and sales volume is analyzed by actual time series big data from clothing company 'A'. As a results, the significant analysis results were obtained by analyzing the sales volume of various product categories via proposed PBDA.

References

[1] M. Vojnovic, F. Xu, J. Zhou, "Sampling based range partition methods for big data analytics," *Microsoft Res.*, Microsoft Corp., Redmond, WA, USA, Tech. Rep. MSR-TR-2012-18, 2012.

[2] L. Wu, R. J. Barker, M. A. Kim, K.A. Ross, "Navigating big data with high-throughput, energy-efficient data partitioning," *ISCA '13: Proceedings of the 40th Annual International Symposium on Computer Architecture*, 249-260, 2013.

[3] K. Saritha, S. Abraham, "Prediction with partitioning: Big data analytics using regression techniques," *In Networks Advances in Computational Technologies (NetACT)*, 208-214, 2017.

[4] S. Salloum, J. Z. Huang, Y. L. He, "Random sample partition: a distributed data model for big data analysis," *IEEE Transactions on Industrial Informatics*, vol. 15, 5846-5854, 2019.

[5] S. Jain, J. Bruniaux, X. Zeng, P. Bruniaux, "Big data in fashion industry," *IOP Conf., Mater. Sci. Eng.*, 1-6, 2017.

[6] D. Øivind, T. Stenhei, "Big Data viewed through the lens of management fashion theory," *Cogent Business Management* vol. 3, 1-19, 2016.

Factors Influencing the Interoperability of Building Information Modeling in Architecture, Engineering and Construction Industry

Hafiz Muhammad Faisal Shehzad *[1], Roliana Binti Ibrahim [2], Ahmad Fadhil Yusof [3], Khairul Anwar Mohamed khaidzir [4], Israr Ghani [5] and Abba Kyari Buba[6]
[1, 2, 3, 5, 6] School of Computing, University Technology Malaysia, Johor Bahru, Malaysia.
[4] Department of Architecture, Faculty of Built Environment, University Technology Malaysia, Johor Bahru, Malaysia.
[1]Department of Computer Science & IT, University of Sargodha, Punjab, Pakistan
[E-mails: muhammad.faisal@uos.edu.pk [1], roliana@utm.my [2], ahmadfadhil@utm.my [3], b-anwar@utm.my [4], israrghani@outlook.com [5], bkoriyell@gmail.com [6]]
*Corresponding author: Hafiz Muhammad Faisal Shehzad

Abstract

Digital technology is a key component of collaboration between the various stakeholders involved in construction projects. Through introducing Building Information Modelling (BIM), technology revolutionized the Architecture Engineering and construction industry (AEC) working practices. Despite BIM's numerous benefits, its adoption in AEC is susceptible to confrontation. Interoperability is the main obstacle often cited for BIM adoption. However, the factors influencing interoperability got limited attention in the literature. Therefore this research article aims to identify the influencing interoperability factors in AEC for BIM adoption. For this purpose, this study performs a content analysis of existing BIM standards and guidelines published in different countries. The identified factors will help AEC professionals to improve interoperability activities, coordination, and mapping of the organizational process to make the collaborative project delivery more effective and well informed. Also, the findings will assist policymakers in developing a roadmap to overcome the obstacles to interoperability.

***Keywords*:** BIM Interoperability, Building Information Modelling (BIM), Interoperability factors, Collaborative project delivery

1. Introduction

The building industry includes many construction stakeholders such as architects, project managers, quantity inspections, manufacturers, plant managers, mechanical engineers, and plumbing engineers. Due to the participation of stakeholders from diverse disciplines, construction is a complex activity. Delay in the project, poorly design documents, cost over-runs [1], lack of collaboration [2], and interoperability of construction software are the major challenges in the construction industry. The recent paradigm shift from traditional delivery processes of AEC is Building Information Modelling (BIM), which can increase productivity, enhanced collaboration,

increase quality, and reduce cost, and project delivery on time. BIM allows the management of digital data models throughout the entire building life cycle, from design to the actual handover of the project to stakeholders.

Although BIM innovation has distinct benefits, its utilization is considerably slower[3], [4]. To facilitate the entire construction activity, BIM uses the shared digital representation of built environments [5], [6]. Several factors affect the adoption of BIM, such as resistance to change, the readiness of the industry [7], [8], knowledge and education [2], [9], and legal constraints [3], [10]. Interoperability is the most cited construct that affects collaborative project delivery. Therefore it is necessary to understand interoperability drivers and factors. However, there is limited focus on interoperability factors in existing BIM research. Therefore the purpose of this study is to find factors effect on interoperability adoption in BIM-enabled construction projects. The rest of the research article provides the related works in section 2, methodology in section 3, and results and conclusion in section 4 and section 5.

2. Related Works

BIM Interoperability means the ability of BIM software to collect, process, exchange and share the same data using the same protocols, and common file formats [11], [12]. BIM's interoperability improves coordination and cooperation between stakeholders in the scattered construction industry [13]. The existing studies on the BIM domain suggest ontologies for mapping cross-discipline data [14], and objects integrations in different disciplines [15]. Similarly few studies focus on interoperability assessment using Industry Foundation Class (IFC) [14] and enhancement of IFC schema to increase data interoperability [16]. Also, the use of a web semantic web can enhance data interoperability [17]. Interoperability is not only concerned with data but also organizational compatibility influence on interoperability [18].

The interoperability issues faced by structural engineers are due to different IFC standards and file formats [19]. Pishdad-Bozorgi et al. [20] suggest a well-implemented plan of interoperability for data exchange between BIM tools enables FM-enabled BIM to be achieved. Addressing the challenges of interoperability requires a collective understanding between construction professionals and environmental flexibility needs. [21]. There are limited studies that stress to find interoperability factors [22]–[24]. Based on the existing literature review, it is evident that existing studies discuss interoperability in general and factors of interoperability got limited attention in the existing literature. Therefore, the novelty of the research study is to identify the most common factors influencing interoperability in AEC based on content analysis of exiting BIM guidelines.

3. Methodology

In this research study, the methodology adopted is based on two steps. The first step is a comprehensive literature review of interoperability in various fields, such as software engineering , e-government and information systems to explore existing interoperability models, such as the European framework for interoperability [25]. The objective of this step is to align the concept of interoperability to the AEC knowledge field. Secondly, existing literature examines the extent to which interoperability concepts are diffused in the AEC research field is reviewed. BIM guidelines, standards, and contract documents have been designed in the past by several organizations, such as government, owner organizations, and AEC firms to facilitate the implementation of BIM's projects. An analysis of the qualitative contents of existing BIM standards and guidelines published in various countries is analyzed by following the guidelines [26], [27]. 81 guidelines and standards published by several associated organizations are listed by Building Smart [28]. A total of 15 documents have been considered for the study as given in **Table 1.** The choice of these fifteen documents is based on the AEC industry's chronological precedence, score, and scope of advertising. These documents include standards at the government level, the owner's manual, and guidance published by academic institutes. The main topic of analysis carried out from these documents is the interoperability construct. Interoperability, data integration, data management, open architectures, and data models are the keywords looked for from these standards. For the commonly recurring themes,

both the content and thematic analysis techniques were practised. The analysis process involved two investigators. In order to minimise bias, the analysis results of both the investigators were compared and the census was developed.

4. Results and Analysis

The topics reviewed in these guidelines include interoperability approaches, legal factors, organizational BIM vision, collaboration environments, and the use of open data standards.To solve data interoperability problems, use of IFC and COBIE data standards is emphasized in most of the guidelines. Open BIM is also proposed to maximize data interoperability. As for organizational interoperability, only BPC, NATSPECNB, VABG, and ACOE concentrate on the need for organizational interoperability. On the other hand, limited details regarding organizational interoperability is found in BLACCD and CAS. In current guidelines, semantic and legal interoperability has received very limited attention.The OSU, VABG covers limited aspects of semantic interoperability and few details of legal interoperability are described by BPC. The factors contributing to interoperability as set out in **Table. 2** are also examined in the guidelines. In the current BIM guidelines, interoperability factors such as data security and data integration have received a great deal of attention. In the BIM guidelines, however, BIM complexity is not much discussed. Similarly, in BSI, CAS, NBIMSUS, VABG, and SBM , BIM compatibility is discussed. Other studies have demonstrated limited concentration on compatibility. In BIM guidelines, financial constraints have received limited attention.

The common definition, creation of data dictionaries, workflow mapping, and development of exchange standards are the focal points of semantic interoperability as provided in ACOE, COBIM, and IUA. Contraty to former suggestions the other 12 guidelines stressed the development of exchange standards.Similarly, the common definition and the creation of data dictionaries are also suggested by standards to prevent issues of misinterpretation during data transfer. To attain semantic interoperability BSI, CAS, and BIGUSE, stress on mapping workflows.

5. Conclusions

BIM Interoperability is the ability to share information and understandably share data in two or more organizational units or networks. In a narrow sense, interoperability is only considered as technological issues and studies aim to improve the applications for interoperability, however, some studies suggest that interoperability is being explored in various directions. Therefore this research article aims to identify the influencing interoperability factors in AEC for BIM adoption. For this purpose, this study performs a content analysis of existing BIM standards and guidelines published in different countries. The findings will assist policymakers in developing a roadmap to overcome the obstacles to interoperability. Focusing on the identified factors will help AEC professionals to improve interoperability activities, coordination, and mapping of the organizational processes to make the collaborative project delivery more effective and well informed. The limitations of the study are the selection of the limited number of BIM guidelines therefore future studies should consider enhanced sample documents. Future guidance for the work is the validation of the findings by collecting data, which will open a new discussion on interoperability between the BIM domain and AEC stakeholders.

5. References

[1] H. M. Faisal Shehzad, R. Binti Ibrahim, A. F. Yusof, K. A. Mohamed khaidzir, S. Shawkat, and S. Ahmad, “Recent developments of BIM adoption based on categorization, identification and factors: a systematic literature review,” *Int. J. Constr. Manag.*, vol. 0, no. 0, pp. 1–13, 2020.

[2] H. M. F. Shehzad, R. B. Ibrahim, A. F. Yusof, and K. A. M. Khaidzir, “Building Information Modeling: Factors Affecting the Adoption in the AEC Industry,” in *2019 6th International Conference on Research and Innovation in Information Systems (ICRIIS)*, 2019, pp. 1–6.

Table. 1. BIM Standards and Guidelines

Acronym	Country	Standard Name	Source
GSABIM	US	GSA BIM Guide	The National 3D-4D-BIM Program [29]
ACOE	US	Official Manual For BIM projects	Army Corp [30]
COBIM	Finland	COBIM (Common BIM Requirements)	Senate Properties [31]
IUA	US	BIM Guidelines &Standards for Architects, Engineers, and Contractors	Indiana University [32]
BIGUSC	US	Building Information Guidelines	USC [33]
NATSPECNB	Australia	NATSPEC National BIM Guide v1.0	NATSPEC [34]
CAS	Canada	AEC (CAN) BIM Protocol	Canadian Standards [35]
BPC	Singapore	BIM Particular Conditions version 2.0	Building and Construction Authority [36]
NBIMSUS	US	National BIM Standard - United States - Version 3	BuildingSMART [37]
GIT	US	Georgia Tech BIM Requirements & Guidelines for Architects, Engineers, and Contractors	Georgia Institute of Technology [38]
VABG	US	BIM Manual v 2.2	Affairs [39]
BLACCD	US	LACCD BIM standards	Build Laccd [40]
SBM	Norway	Statsbygg BIM manual 1.2.1	Statsbygg [41]
BSI	UK	Little Book of BIM	The British Standards [42]
OSU	US	Building Information Modelling Project Delivery Standards	Ohio state university [43]

Table. 2. Interoperability Factors

Interoperability Factors	BSI	CAS	BPC	NATSPECNB	OSU	VABG	ACOE	GSABIM	SBM	COBIM	IUA	BIGUSC	GIT	BLACCD
Contractual Environments		*	*			*		*						
Intellectual Property		*	*	*	*	*		*					*	*
Regulatory support			*											*
Insurance framework											*			
Work Flow Mapping	*	*	*	*		*						*		*
Common definitions	*	*		*	*	*		*	*			*	*	*
Exchange standard	*	*	*	*	*	*		*	*			*	*	*
Financial Constraints											*			
Data security		*	*	*		*		*					*	*
Compatibility	*	*				*			*					
Complexity														
Data integration	*	*	*	*		*			*			*	*	*
Data Dictionaries				*		*			*			*	*	*

[3] D. Walasek and A. Barszcz, "Analysis of the Adoption Rate of Building Information Modeling [BIM] and its Return on Investment [ROI]," in *Procedia Engineering*, 2017, vol. 172, pp. 1227–1234.

[4] S. Ahmed, "Barriers to Implementation of Building Information Modeling (BIM) to the Construction Industry: A Review," *J. Civ. Eng. Constr.*, vol. 7, no. 2, p. 107, 2018.

[5] ISO, "ISO 29481-1:2016 - Building information models -- Information delivery manual -- Part 1: Methodology and format," 2016. [Online]. Available: https://www.iso.org/standard/60553.html. [Accessed: 17-Jan-2019].

[6] M. M. Khurshid, N. H. Zakaria, A. Rashid, H. Muhammad, and F. Shehzad, "Modeling of Open Government Data for Public Sector Organizations Using the Potential Theories and Determinants—A Systematic Review," *informatics*, 2020.

[7] W. I. Enegbuma, G. U. Aliagha, K. N. Ali, and Y. Y. Badiru, "Confirmatory strategic information technology implementation for building information modelling adoption model," *J. Constr. Dev. Ctries.*, vol. 21, no. 2, pp. 113–129, 2016.

[8] M. Btoush and A. T. Haron, "Understanding BIM Adoption in the AEC Industry: The Case of Jordan," in *IOP Conference Series: Materials Science and Engineering*, 2017, vol. 271, no. 1.

[9] J. Zhang, K. Schmidt, and H. Li, "BIM and sustainability education: Incorporating instructional needs into curriculum planning in CEM programs accredited by ACCE," *Sustain.*, vol. 8, no. 6, Jun. 2016.

[10] A. L. Ahmed, J. P. Kawalek, and M. Kassem, "A comprehensive identification and categorisation of drivers, factors, and determinants for BIM adoption: A systematic literature review," in *Computing in Civil Engineering 2017*, 2017, pp. 220–227.

[11] A. Grilo and R. Jardim-Goncalves, "Value proposition on interoperability of BIM and collaborative working environments," *Autom. Constr.*, vol. 19, no. 5, pp. 522–530, 2010.

[12] buildingSMART, "Technical Vision - buildingSMART," 2016. [Online]. Available: https://www.buildingsmart.org/standards/technical-vision/. [Accessed: 17-Jan-2019].

[13] M. F. Muller, A. Garbers, F. Esmanioto, N. Huber, E. R. Loures, and O. Canciglieri, "Data interoperability assessment though IFC for BIM in structural design–a five-year gap analysis," *J. Civ. Eng. Manag.*, vol. 23, no. 7, pp. 943–954, 2017.

[14] Y.-C. Lee, C. M. Eastman, and J.-K. Lee, "Validations for ensuring the interoperability of data exchange of a building information model," *Autom. Constr.*, vol. 58, pp. 176–195, 2015.

[15] K. Kim, H. Kim, W. Kim, C. Kim, J. Kim, and J. Yu, "Integration of ifc objects and facility management work information using Semantic Web," *Autom. Constr.*, vol. 87, pp. 173–187, 2018.

[16] P. Matějka, V. Kosina, A. Tomek, R. Tomek, V. Berka, and D. Šulc, "The Integration of BIM in Later Project Life Cycle Phases in Unprepared Environment from FM Perspective," *Procedia Eng.*, vol. 164, pp. 550–557, Jan. 2016.

[17] P. Pauwels, S. Zhang, and Y.-C. Lee, "Semantic web technologies in AEC industry: A literature overview," *Autom. Constr.*, vol. 73, pp. 145–165, 2017.

[18] J. Zhang, Q. Liu, Z. Hu, J. Lin, and F. Yu, "A multi-server information-sharing environment for cross-party collaboration on a private cloud," *Autom. Constr.*, vol. 81, pp. 180–195, 2017.

[19] Z.-Q. Liu, F. Zhang, and J. Zhang, "The building information modeling and its use for data transformation in the structural design stage," *J. Appl. Sci. Eng.*, vol. 19, no. 3, pp. 273–284, 2016.

[20] P. Pishdad-Bozorgi, X. Gao, C. Eastman, and A. P. Self, "Planning and developing facility management-enabled building information model (FM-enabled BIM)," *Autom. Constr.*, vol. 87, no. February

2017, pp. 22–38, 2018.

[21] Y. Arayici, T. Fernando, V. Munoz, and M. Bassanino, "Interoperability specification development for integrated BIM use in performance based design," *Autom. Constr.*, vol. 85, no. November 2017, pp. 167–181, 2018.

[22] E. of I. in the A. I. A. Poirier, D. Forgues, and S. Staub-French, "Dimensions of Interoperability in the AEC Industry," in *Construction Research Congress 2014*, 2014, pp. 1987–1996.

[23] H. Xu, J. Feng, and S. Li, "Users-orientated evaluation of building information model in the Chinese construction industry," *Autom. Constr.*, vol. 39, pp. 32–46, 2014.

[24] M. F. Muller, A. Garbers, F. Esmanioto, N. Huber, E. R. Loures, and O. Canciglieri, "Data interoperability assessment though IFC for BIM in structural design – a five-year gap analysis," *J. Civ. Eng. Manag.*, vol. 23, no. 7, pp. 943–954, 2017.

[25] C. European, "New European Interoperability Framework," *European commision*, 2017. .

[26] H.-F. Hsieh and S. E. Shannon, "Three Approaches to Qualitative Content Analysis," *Qual. Health Res.*, vol. 15, no. 9, pp. 1277–1288, 2005.

[27] P. Mayring, "Qualitative Content Analysis," *FORUM Qual. Soc. Res.*, vol. 1, no. June, p. 20, 2000.

[28] BuildingSmart, "'BIM Guide Project,'" *BuildingSmart*, 2015. [Online]. Available: http://bimguides.vtreem.com/bin/view/BIMGuides/Guidelines. [Accessed: 25-Sep-2019].

[29] O. of the The National 3D-4D-BIM Program and P. B. S. Chief Architect, "GSA Building Information Modeling Guide Series 01 – Overview," 2007.

[30] Army Corp. of Engineers, "New York District US Army Corp. of Engineers official manual for building information modeling projects," 2009.

[31] Senate Properties, "Common BIM Requirements," 2012.

[32] Indiana University Architect's Office, "Indiana University Building Information Modeling (BIM) Guidelines and Standards for Architects, Engineers, and Contractors," 2012.

[33] C. C. D. and F. M. S. USC, "University of Southern California Building Information Modeling (BIM) Guidelines version 1.6 For Design Bid Build Contracts," 2012.

[34] NATSPEC Construction Information, "NATSPEC National BIM Guide v1.0," 2016.

[35] I. Canadian and B. I. M. Standards, "AEC (CAN) BIM Protocol," 2014.

[36] Building and Construction Authority, "BIM Particular Conditions," 2015.

[37] buildingSMART alliance, "NBIMS-US-3: 5.2 Minimum BIM," *Natl. BIM Stand. - United States - Version 3*, p. 13, 2015.

[38] Georgia Institute of Technology, "Georgia Tech BIM Requirements & Guidelines for Architects, Engineers and Contractors," 2016.

[39] U. D. of V. Affairs, "BIM Manual v 2.2," 2017.

[40] Build Laccd, "LACCD Building Information Modeling Standards," 2017.

[41] B. Statsbygg, "STATSBYGG BIM-MANUAL 2.0 (SBM2)," 2019.

[42] I. The British Standards, "Little Book of BIM," France, 2019.

[43] The Ohio state university, "Building Information Modeling (BIM) Project Delivery Standards," 2019.

An Enhanced String Algorithm Prioritizing Technique

M. Irsyad Kamil R.[1], Dayang N.A. Jawawi[2], Rosbi Mamat[3]
[1, 2]School of Computing
[3]School of Electrical Engineering,
Faculty of Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia,
[e-mail: mirsyad6@live.utm.my, dayang@utm.my, rosbi@utm.my]

Abstract

In modern times Information Technology (IT) has provided users with various solution. With the continuous evolution of software systems, test suites often grow very large. Rerunning all test cases may be impractical in regression testing; thus, reordering all test case is one of the ways to generate satisfactory results. The proposed paper is determined in producing a better approach in delivering regression testing with a high fault detection rate. It presents an enhancement on string algorithm where it will be applied on measuring the similarity or distance between test cases in a test suite. This method will assist prioritisation technique in regression testing. There are five string algorithms tested in the experiment, including the proposed enhancement. The Jaro-Winkler similarity was selected to be enhanced due to its Average Percentage Rate Detection (APFD) value in the initial experiment. On the early results, the APFD rate for Jaro-Winkler is 87.64% on average which is lower than other algorithms. So, we believed that an improvement could be made on the equation to increase the APFD rate. We propose an enhancement on the algorithm by merging Jaccard Index equation. We compared our proposed approach to the naive approach. Our results prove a significant improvement in the efficiency where the enhancement made a 5.63% increase rate and 93.27% on average.

***Keywords*:** Jaro-Winkler, Similarity-based, Test Case Prioritisation

1. Introduction

Software testing is essential to ensure that the software developed is error-free. There have been many kinds of testing these days, and each test aims to eliminate errors from a particular point of view. Regression testing, for example, is a test where the quality of the system is ensured after new features have been added to the system. Regression testing is widely prevalent among testers, and many existing ways can be used to carry out the test. The execution order of the test case is critical, and hence more notable test cases should be executed as early as possible in a test set so that faults can be detected faster [1]. Calculating similarity among the test cases is essential as it can contribute to reordering the test execution order.

The method of assessing test execution order is called test case prioritisation (TCP). To find faults, it prioritises test cases with the highest possibility of executing them. Many TCP algorithms were proposed to guide the priority settings for various types of test cases, such as codes-coverage prioritisation [2] and risk-based prioritisation [3]. On the other hand, the number of studies using similarity/distance information from test cases has recently increased to guide the execution [4], meaning that it is advisable to implement or explore more on similarity-based approach.

This research was supported by a research grant from Universiti Teknologi Malaysia (UTM), UTM-TDR Grant Vot No. 06G23 and Ministry of Higher Education (MOHE), Grant Vot No. 5F117

Most recently, a similarity-based test case prioritisation techniques using coverage-based have been developed [5]. The idea of the developed has proven to be useful as they managed to increase the fault detection rate. Using the concept of combining similarity strategy with more than one coverage criteria have led to the idea of combining two or more string matrices in finding the similarity. Moreover, the idea is not unusual since a few years back, a similarity-based test case prioritisation technique with enhanced string matrices have been developed [6]. Since the similarity-based test case prioritisation (STCP) is a cost-effective prioritisation technique [7], it is most likely to implement the approach in our experiment.

In this paper, we proposed an enhance method where we combine two algorithms named Jaro-Winkler and Jaccard Index. Furthermore, the measurement using Jaccard Index (JI), Cosine Similarity (CS), Levenshtein Distance (LD) and Jaro-Winkler (JW) metric, on the similarity-based prioritisation algorithm were taken and compared with the proposed algorithm. The experiment dataset is taken from Seimens dataset, and all codes for experimenting are written in the JavaScript language.

The rest of this paper is organised as follows: Section 2 describes related works. Experimental setup and results are presented in Section 3 and Section 4. Section 5 concludes our works and explained future works.

2. Related Work

The similarity-based approach is commonly applied to TCP and TCS. Each approach that was implemented have a different set of goals; it can be cost-effective [1], high fault detection rate [5] and less time [4]. Regression testing is a very time-consuming and expensive activity [8]. Hence, it justifies the notion behind these approaches. Similarity-based can make testing to be time and cost-effective. Several existing studies had applied the same approach in solving their challenges, and these articles are similar to ours and worth mentioning.

In the work of [6], the authors enhanced string distance formula by combining Jaro-Winkler and Hamming Distance equation. Prioritisation algorithm was also improved by removing and substitute a few lines of codes in All-yes Config. The purpose of this enhancement was to increase the effectiveness of similarity-based prioritisation technique on Software Product Line (SPL) and based on the outcome of the research, the effectiveness of SPL testing process has increased in terms of APFD scores, execution time and rate of fault detection.

Besides that, in the work of [5], the authors implement similarity-based prioritisation technique with an aim, to test any software with minimal time and maximum efficiency in terms of fault coverage rate. Some element from code-coverage-based were added into the experiment process. The experiment did not only look into similarity value but also coverage among the test suite. The idea of multi-criteria is astonishing as the results showed their technique are practicable and quite significant compared to traditional approaches.

Meanwhile, the work by [9] presented a hybrid weight-based and string distance using Particle Swarm Optimization (PSO) in prioritisation. The hybridisation was capable of improving APFD values compared to non-hybrid string distance showed by the outcome.

Hybridisation is quite common within these three years, and two articles applied the process [5][9]. The method that they used, hybridisation, are quite similar to the enhancement made by them [6]. All of them managed to combine and merge certain equation and process which motivate us to apply their methods in our research.

3. The Enhanced String Algorithm

In this section, the process of enhancing the string algorithm is described.

On the early stage of our experiment, we identify our research domain, objectives and scope. Once the research foundation is formulated, the following research components were thoroughly reviewed. The components include prioritisation algorithm and string algorithm. For the prioritisation algorithm, there are three algorithms: Local Maximum Distance (LMD), Global Maximum Distance (GMD) and All-yes Config. Meanwhile, for selection algorithm, there are four algorithms: Jaro-Winkler, Jaccard

Index, Cosine Similarity and Levenshtein Distance. However, only LMD will be applied in this experiment because it can quickly find faults in large models compared to other algorithms [10]. A summary of the components involved in this experiment is shown in **Table 1**.

Table 1. A summary of components involved

Components	Items
Prioritisation algorithm	Local maximum distance
Selection algorithm	Jaro-Winkler, Jaccard index, Cosine similarity, Levenshtein distance

On the next phase, we execute the experiment based on the framework that we prepared earlier in the first phase. Siemens dataset was used in the experiment. Many researchers have used this test suite as their experiment study, and this is one of the many papers existed [9], [11], [12]. An overview of the dataset used for this experiment is shown in **Table 2**.

Table 2. An overview of the dataset used

Dataset name	Programming language	Test cases	Fault matrix
Tcas	C	1608 test cases	Have different fault matrix
j-tcas	Java		
Cs-tcas	C#		

The experiment calculates the similarity of test cases by using all string algorithms that we mentioned earlier. Then, in the evaluation phase, the prioritised test suites are evaluated by calculating their APFD rate based on fault matrix of the program. The outcome from the experiment will able to determine which algorithm contribute to the highest rate of fault detection.

There is two types of experiment: initial experiment and final experiment. Both experiments undergo the same process. However, results from the initial experiment will determine which algorithm undergoes enhancement. In our case, Jaro-Winkler was selected, and we chose the Jaccard Index as a factor to merge with because both algorithms are character-based. Jaro-Winkler's origin is formed in 1999 by William E. Winkler [13]. The equation was an enhanced version of Jaro Similarity, which was created a year before the introduction of Jaro-Winkler. The equation can be split into three factors: length of string, matching characters and transposition. Jaccard index formation is similar to Jaro-Winkler; it also has matching characters and transposition elements in its formula. Thus, the merge is possible due to the similarity of factors within each equation. With the formation of the new enhanced Jaro-Winkler, we believed that it would contribute to our experiment result. An illustration on merging both formulae are shown in **Fig. 1**.

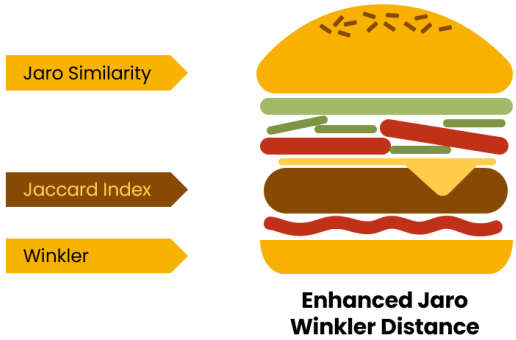


Fig. 1. Enhanced Jaro Winkler illustration

4. Experiment Result

The overall phase of the experiment is where the authors concluded the studies and suggest some future works that could help in increasing the effectiveness of regression testing. The whole process of the experiment can be seen below in **Fig. 2**.

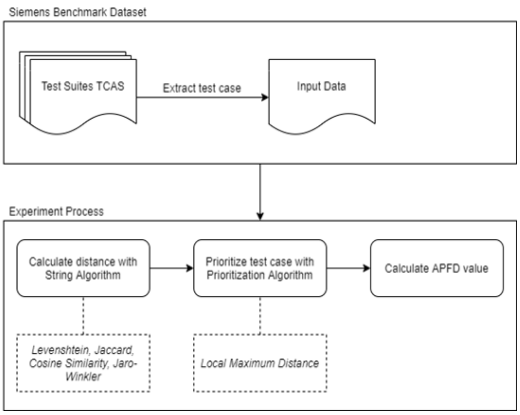


Fig. 2. Experiment process

4.1 Initial Experiment

In this initial experiment, we start experimenting by calculating the similarity between the test cases in the dataset with all string algorithms mentioned before. After calculating the similarity, the experiment will undergo prioritisation using LMD. Then, we calculate the values of APFD for each test execution order produced by each string algorithm. The result of the APFD is recorded in **Table 3**.

Table 3. Result APFD

String algorithm	APFD Rate (%)		
	tcas	jtcas	cstcas
Levenshtein	64.22	93.89	88.72
Jaccard	91.93	92.50	92.75
Cosine similarity	92.70	93.65	93.21
Jaro-winkler	87.31	88.04	87.57

Results for **Table 3** clearly shows an indication where we should apply an enhancement on Jaro-Winkler in making sure the result reach to at least 90%. The APFD rate values in **Table 3** showed Cosine Similarity is the best in finding faults. The results in Table 3 have led us to the next exploration of the enhancement of Jaro-Winkler.

2.1 Final Experiment

In this final experiment, the process is similar to the initial experiment, and the only difference is the algorithm. In this experiment, the proposed algorithm is executed.

We proposed a merge equation of Jaro-Winkler with Jaccard because both of there equation are character-based equation. We did try to merge other equation as well, but the result did not turn out to be as good as the enhanced Jaro-Winkler. The enhanced equation is shown below in **Fig. 2**.

$$d_w = d_j + \ell p(1 - d_j)$$

$$d_j = \frac{1}{3}\left(\frac{|A \cap B|}{|S_1|} + \frac{|A \cap B|}{|S_2|} + \frac{||A \cap B| - (|A \cup B| - |A \cap B|)|}{|A \cap B|}\right)$$

Fig. 3. An enhanced equation of Jaro-Winkler

Then, we recalculate the APFD values, but this time, we include our enhanced version of Jaro-Winkler. The result of the APFD values is shown in **Table 4**.

Table 4. APFD values with EJW

String algorithm	APFD Rate (%)		
	tcas	jtcas	cstcas
Levenshtein	64.22	93.89	88.72
Jaccard	91.93	92.50	92.75
Cosine similarity	92.70	93.65	93.21
Jaro-winkler	87.31	88.04	87.57
Enhanced Jaro-Winkler	92.22	94.14	93.45

The result, in **Table 4** shows that the enhanced string metric yields the highest APFD in tcas, jtcas, and cstcas with scores of 90.22%, 92.14%, and 91.45% respectively. These findings prove the author's suggestions that the extension Jaro-Winkler when combined with the Jaccard Index, can increase the APFD scores, which helps in revealing defects the earliest.

5. Conclusions & Future Work

Throughout this paper, we proposed an enhanced string algorithm which helps in calculating the similarities for prioritisation technique. This approach assists in generating a better similarity and more diverse which help the prioritisation algorithm in reordering test cases. The experiment was performed on Siemens dataset. In terms of findings, the proposed enhanced string technique yields better result in terms of APFD. Enhanced Jaro-Winkler (EJW) produced highest APFD scores of 92.22% for tcas, 93.45% for schedule, and second-highest 94.14% for jtcas, overall, EJW performed among the best, scoring scored above 91%.

For future works, we will be applying similarity-based on selection process while prioritising the test suite order. Besides that, we are also always wanted to add multiple criteria in our research since the idea has been proven to be effective much recently [14][15].

Acknowledgement

The authors wish to acknowledge Universiti Teknologi Malaysia for UTM-TDR Grant Vot No. 06G23 and Ministry of Higher Education (MOHE), Grant Vot No. 5F117, which have made this research possible. The authors would also like to express their warmest appreciation on

continuous support and inputs provided by members of Embedded Real-Time Software Engineering Laboratory.

References

[1] R. Huang, Y. Zhou, W. Zong, D. Towey, and J. Chen, "An Empirical Comparison of Similarity Measures for Abstract Test Case Prioritization," *Proc. - Int. Comput. Softw. Appl. Conf.*, vol. 1, pp. 3–12, 2017, doi: 10.1109/COMPSAC.2017.271.

[2] I. Alazzam and K. M. O. Nahar, "Combined source code approach for test case prioritisation," *ACM Int. Conf. Proceeding Ser.*, pp. 12–15, 2018, doi: 10.1145/3209914.3209936.

[3] C. Hettiarachchi, H. Do, and B. Choi, "Risk-based test case prioritisation using a fuzzy expert system," *Inf. Softw. Technol.*, vol. 69, pp. 1–15, 2016, doi: 10.1016/j.infsof.2015.08.008.

[4] F. G. de Oliveira Neto, A. Ahmad, O. Leifler, K. Sandahl, and E. Enoiu, "Improving continuous integration with similarity-based test case selection," in *AST*, 2018, pp. 39–45, doi: 10.1145/3194733.3194744.

[5] S. Singh, P. Ranjan, and R. Shree, "Improving fault detection rate using similarity-based test case prioritisation in regression testing," *Int. J. Innov. Technol. Explor. Eng.*, vol. 9, no. 1, pp. 1874–1880, 2019, doi: 10.35940/ijitee.A5191.119119.

[6] M. Sahak, "Similarity Distance Measure and Prioritization," vol. 1, no. 1, pp. 57–75, 2019.

[7] M. Al-Hajjaji, T. Thüm, J. Meinicke, M. Lochau, and G. Saake, "Similarity-based prioritisation in software product-line testing," in *Proceedings of the 18th International Software Product Line Conference on - SPLC '14*, 2014, pp. 197–206, doi: 10.1145/2648511.2648532.

[8] R. Wang, S. Jiang, and D. Chen, "Similarity-based regression test case prioritisation," *Proc. Int. Conf. Softw. Eng. Knowl. Eng. SEKE*, vol. 2015-Janua, no. 1, pp. 358–363, 2015, doi: 10.18293/SEKE2015-115.

[9] M. Khatibsyarbini, M. A. Isa, and D. N. A. Jawawi, "A hybrid weight-based and string distances using particle swarm optimisation for prioritising test cases," *J. Theor. Appl. Inf. Technol.*, vol. 95, no. 12, pp. 2723–2732, 2017.

[10] B. Lewis, I. Smith, M. Fowler, and J. Licato, "The robot mafia: A test environment for deceptive robots," *28th Mod. Artif. Intell. Cogn. Sci. Conf. MAICS 2017*, pp. 189–190, 2017, doi: 10.1145/1235.

[11] B. Jiang, Z. Zhang, W. K. Chan, T. H. Tse, and T. Y. Chen, "How well does test case prioritisation integrate with statistical fault localisation?," *Inf. Softw. Technol.*, vol. 54, no. 7, pp. 739–758, 2012, doi: 10.1016/j.infsof.2012.01.006.

[12] Y. Ledru, A. Petrenko, S. Boroday, and N. Mandran, "Prioritising test cases with string distances," *Autom. Softw. Eng.*, vol. 19, no. 1, pp. 65–95, 2012, doi: 10.1007/s10515-011-0093-0.

[13] W. E. Winkler, "The State of Record Linkage and Current Research Problems," *Stat. Res. Div. US Census Bur.*, pp. 1–15, 1999, doi: 10.1.1.39.4336.

[14] S. Singh and R. Shree, "SPECIAL ISSUE REDSET 2016 OF CSIT A combined approach to optimise the test suite size in regression testing," *CSI Trans. ICT*, 2016, doi: 10.1007/s40012-016-0098-8.

[15] C. Sharma, "Mechanism for Identification of Duplicate Test Cases," pp. 0–4, 2014.

Finding ways to improve student satisfaction with online-based team activities

Sangwoo Hahm, Marvin Salganov, Kundong Choi
Department of Business Administration, Semyung University, Chungbuk, South Korea
Bluline, FL, USA 07072
Purchasing Team, Kwang Dong Pharmaceutical Co., Ltd. Seoul, 06650, Korea
[e-mail: thehahm@semyung.ac.kr, marvin@bluline.com, d881414@hanmail.net] *
Corresponding author : Kundong Choi

Abstract

Most universities are striving to develop professors' technical skills in online lectures, communicate with students, and improve lecture satisfaction in the post-COVID-19. These efforts will improve students' satisfaction with online education. Professors are also making various efforts to improve the quality of online education. The development of various new educational methods will improve learners' satisfaction. Among various teaching methods, this study focuses on the characteristics of online education based on team activities. Through team activities, students can gain various experiences and gain practical knowledge. In particular, from the perspective of business administration, these activities are closely related to student performance. Every basic unit of work today is a team. Online-based team activities have advantages such as running a virtual team, communication, minimizing emotional conflicts, and exchanging fast and massive amounts of information. Hence, if students gain educational experience in online-based team activities, their performance will improve in the future. As a way to activate these online-based team activities, this study proposes 1) the use of flip learning and action learning 2) a platform that supports real-time discussions and 3) the importance of student participation and motivation. .

Keywords*:* Online-based education, online-based team activities, student satisfaction

1. Introduction

Online-based education will be activated in the post-COVID-19. Most universities are striving to develop professors' technical skills in online lectures, communicate with students, and improve lecture satisfaction. These efforts focus on the relationship between professors and students. However, the relationship between students can also affect educational outcomes. Among certain courses in universities, there are cases where students' team activities are required. Team activities can bring about a variety of educational effects such as creating new values, developing teamwork, improving skills of communication, building a cooperative attitude, and case studies.

Students will experience team activities using online media. In order to increase the effectiveness of online-based team activities, new and diverse efforts will be required at the level of schools, professors, and students that make up the education system. First, university should endeavor to provide an easy-to-use system, student-centered operation, actively resolve errors and inquiries, and accept the opinions of students and professors. Second, professors should inform students of the clear goals and methods of team activities, provide feedback and interest in team activities, and conduct a fair evaluation. Third, students will need to improve their ability to use the system for

online-based team activities, eliminate free rides, and participate voluntarily.

Team activity is a teaching method that entails a variety of strengths. If schools and professors support online-based team activities, and students participate, more successful online-based team activities will occur and educational outcomes will be improved. These results will enhance students' perceived satisfaction with university education. Furthermore, students will be able to develop competency in online-based team activities through these experiences. Companies will prefer students with online-based competencies. They will be able to have online-based communication skills when they get a job at a company or work as a team. In addition, the performance of an online-based team of these members could also be improved.

2. Literature review

2.1 Flipped learning and action learning

Flipped learning refers to education in a reverse-paced teaching method, different from the form in which teachers provide information and students learn in traditional ways. Reverse-paced education means a form in which teachers and students share data in advance for education and then use them for discussion or team activities. Flipped learning can increase educatonal effectiveness when students have the will for learning and self-regulated competence. In addition, flipped learning has the advantage of creating various outcomes in a new form of education, enhancing practical learning effects, and enabling strategical usage. In addition, flipped learning can be used more actively in classes through online education [1-3]. Therefore, students will be able to use this form of education for more active team activities after learning by first sharing materials through online-based education.

Action learning refers to a form of education in which students are in charge of actual work and solve the problem with the help of professors, facilitators, or mentors. The basis concept of action learning is based on the learning principle of learning by doing, and generally supports team activities and team learning. Action learning has the advantage of improving general educational outcomes and being applicable in the field in relation to real team activities and practical problem solving [4-6]. Action learning is an experience-oriented education method in which students actually experience and do activities. If online-based action learning is conducted, students can interact in a variety of ways. You will have opportunities to improve your achievements, such as dialogue, solving common team problems, securing diversity, and improving your creativity.

2.2 Platform that supports real-time team activity

Online education entails a variety of situations. In particular, various situations such as communication problems, information sharing issues, and sustained interactions can occur in the curriculum of team activities in online education. Therefore, for online-based team activity education, it is necessary to develop various platforms that can be applied in various situations [7-8]. Such an online education support system needs to have features such as inducing interest of learners, sharing a large amount of information quickly, easier to use, a learner-friendly navigator, and quick response to questions. The outcome of this education can be improved only by supporting online-based team activities through the development of a more learner-friendly system and various platforms.

2.3 Student participation and motivation

Learner's motivation is a very important factor in education. Motivation is closely related to participation and outcomes in education [9-10]. Therefore, the willingness and motivation of learners to participate in online-based team activities are a prerequisite. Effective participation and motivation have a direct impact on educational outcomes [11].

3. Conclusion

Online education will be more active in the future. Most universities and educational institutions recognize the importance of online education, and this method of education will be more encouraged. Online education will develop in various forms such as flip learning, action

learning, online-based information sharing and team activities, as well as simply providing information. Hence, from the school's perspective, online-based team activity education is one of the forms of education that must be performed. In addition, from the perspective of a company, the use of virtual teams and improvement of team work performance are very basic management activities. Today's environment changes rapidly, and companies that need to create new value must utilize this team-type working unit. IT technologies including the Internet are rapidly developing and diversifying. Therefore, students should receive online-based team activity training to gain experience.

This paper presents ways to improve the performance of online-based team activity training. First, the use of learning and action learning, second, a platform that supports inter-discussion, and third, participation and motivation of students are essential factors that have to be premised for online-based team activity education. If these efforts are used to increase competency of students as online team members, they will produce more positive results, such as improving basic performance, exerting synergies, and securing creativity in online-based team work in the future.

References

[1] S. R. Sletten, "Investigating flipped learning: Student self-regulated learning, perceptions, and achievement in an introductory biology course," *Journal of Science Education and Technology*, VOL. 26, NO. 3, pp. 347-358. 2017.

[2] D. C. van Alten, C. Phielix, J. Janssen, and L. Kester, L, "Self-regulated learning support in flipped learning videos enhances learning outcomes," *Computers & Education*, VOL. 158, 104000, 2020.

[3] Y. El Miedany, "Flipped learning," *Rheumatology Teaching*. Springer, Cham, pp. 285-303, 2019

[4] M. Volz-Peacock, B. Carson, and M. Marquardt, M, "Action learning and leadership development," *Advances in Developing Human Resources*, VOL. 18, NO. 3, pp. 318-333. 2016.

[5] A. J. Imperiale, and F. Vanclay, "Experiencing local community resilience in action: Learning from post-disaster communities," Journal of Rural Studies, VOL. 47, pp. 204-219, 2016.

[6] R. Vince, G. Abbey, M. Langenhan, M, and D. Bell, "Finding critical action learning through paradox: The role of action learning in the suppression and stimulation of critical reflection," *Management Learning*, VOL. 49, NO. 1, pp. 86-106, 2018.

[7] A. Bessi, and E. Ferrara, E, "Social bots distort the 2016 US Presidential election online discussion," *First Monday*, VOL. 21, NO. 11-7, 2016.

[8] B. Xu, N. S. Chen, and G. Chen, "Effects of teacher role on student engagement in WeChat-Based online discussion learning," *Computers & Education*, VOL. 157, pp.1-11, 2020.

[9] A. Majnemer, A, "Importance of motivation to children's participation: A motivation to change," *Physical & Occupational Therapy in Pediatrics*, VOL. 31, NO. 1, pp. 1-3. 2011.

[10] S. Bhatti, S. U. Khan, and S. Akbar, "Effect of Molesting Behavior of Male Athletes on Participation and Motivation of Female Athletes at Public Sector Universities of Sindh Province," *SJESR*, VOL. 3, NO. 2, pp. 322-328, 2020.

[11] A. Tasgin, and Y. Tunc, "Effective Participation and Motivation: An Investigation on Secondary School Students," *World Journal of Education*, VOL. 8, NO. 1, pp. 58-74, 2018.

Solving Job Shop Scheduling Problem for Multi-Objective with Nondominated Sorting Genetic Algorithm

Fuladi Shubhendu and Chang Soo Kim
Department of information system, Pukyong national university
Busan, South Korea
[e-mail: shubhendu06@gmail.com , cskim@pknu.ac.kr]
*Corresponding author: Chang-Soo Kim

Abstract

The job shop scheduling is one of the most occurring problems in production planning. N jobs have to be processed on M different machines. To solve this problem there are a few algorithms but, in this paper Non-dominated genetic algorithm being used for solving the job shop scheduling problem. Two objectives are to minimize the total completion time (make-span) And total weighted earliness and tardiness. The architecture of the non-dominated sorting algorithm is similar to a genetic algorithm. Numerical experiments include 10 * 10 example with processing time, machine sequence, and priority and due date.

***Keywords*:** job shop scheduling, Nondominated sorting genetic algorithm, make-span

1. Introduction

The non-dominated sorting algorithm being used for multi-objective scheduling problems. There are two goals, first is to minimize the total completion time (make-span) and the second is to minimize the total weighted early time and delay time (total weighted earliness and tardiness, TWET). Each workpiece will go through 10 processing operations

2. Architecture of NSGA-II

The architecture of NSGA-II is as shown in the figure below. Its architecture is similar to GA. The only major difference steps are Combine parent and offspring population, Non-dominated sorting, calculate crowding-distance, etc. the entire process of NSGA-II will be integrated. For each iteration, the following actions will be performed until the set conditions are reached.

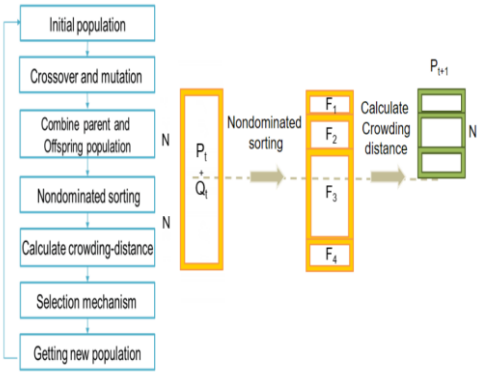


Fig. 1. Architecture of NSGA-II

This work was supported by the Technology Innovation Program (20004205), The development of smart collaboration manufacturing innovation service platform in textile industry by producer-buyer B2B connection funded By MOTIE, Korea.

1. First, there is an initial population (parental generation) P_t which contains N chromosomes, which will produce offspring Q_t after mutation and mating.
2. Due to the elite strategy, the parents and offspring are kept together for selection.
3. Then perform non-overlap sorting to obtain the non-overlap level of each solution (F_1 <level 1>, F_2 <level 2>.....).
4. Finally, select the new N chromosomes as the population for the next iteration. First, select according to the level of non-overlapped. If the following figure occurs, the number of remaining chromosomes to be selected into the new population is less than the next non-overlapped population. The number of chromosomes in the higher level is selected based on the crowded distance, and the larger crowded distance is selected to enter the new population.
5. Finally, a new population P_{t+1} is generated, enter the next iteration, and repeat the above process.

3. Experiment

Suppose there is a job shop scheduling problem with N workpieces and M machines. That chromosome will be composed of N x M genes because each workpiece will only be processed once on each machine. It is processed by M machines, so each workpiece will appear M times in the chromosome. This example is a 10x10 Job shop problem with 10 workpieces and 10 machines. This problem is a multi-objective scheduling problem. There are two goals in total: Minimizing the total completion time (Make-span) and the total weighted early time and delay Time (Total weighted earliness and tardiness, TWET), the workpiece information is shown in the table below. The workpiece information is presented in the processing procedure of the workpiece. Each workpiece will go through 10 processing operations. Three data tables are available here for scheduling one has records of processing time, the second has a record of processing machine sequence and the third is a priority and due date.

3.1 Processing time-

There are 10 jobs (J1, J2, J3, J4, J5, J6, J7, J8, J9, J10) and 10 operations (O1, O2, O3, O4, O5, O6, O7, O8, O9, O10). Every operation has different processing times with respective jobs.

Table 1. Processing time dataset

	A	B	C	D	E	F	G	H	I	J	K
1	order	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10
2	J1	1	2	3	4	5	6	7	8	9	10
3	J2	1	3	5	10	4	2	7	6	8	9
4	J3	2	1	4	3	9	6	8	7	10	5
5	J4	2	3	1	5	7	9	8	4	10	6
6	J5	3	1	2	6	4	5	9	8	10	7
7	J6	3	2	6	4	9	10	1	7	5	8
8	J7	2	1	4	3	7	6	10	9	8	5
9	J8	3	1	2	6	5	7	9	10	8	4
10	J9	1	2	4	6	3	10	7	8	5	9
11	J10	2	1	3	7	9	10	6	4	5	8

3.2 Machine Sequence-

In this table, there is an order of 10 machines. Every operation (O1, O2, O3, O4, O5, O6, O7, O8, O9, O10) goes through every job with the given sequence

Table 2. Machine sequence dataset

	A	B	C	D	E	F	G	H	I	J	K
1	Time	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10
2	J1	29	78	9	36	49	11	62	56	44	21
3	J2	43	90	75	11	69	28	46	46	72	30
4	J3	91	85	39	74	90	10	12	89	45	33
5	J4	81	95	71	99	9	52	85	98	22	43
6	J5	14	6	22	61	26	69	21	49	72	53
7	J6	84	2	52	95	48	72	47	65	6	25
8	J7	46	37	61	13	32	21	32	89	30	55
9	J8	31	86	46	74	32	88	19	48	36	79
10	J9	76	69	76	51	85	11	40	89	26	74
11	J10	85	13	61	7	64	76	47	52	90	45

3.3 Priority and Due date-

In this data set, every job (J1, J2, J3, J4, J5, J6, J7, J8, J9, J10) has its priority and the due date

Table 3. Priority and Due date dataset

	A	B	C
1	order	priority	due date
2	J1	10	919
3	J2	5	785
4	J3	1	907
5	J4	5	849
6	J5	10	887
7	J6	1	783
8	J7	1	670
9	J8	5	861
10	J9	1	801
11	J10	10	896

The raw inputs are needed to carry on other steps like crossover, mutation. Some inputs are population size, crossover rate, size of mutation rate, mutation-selection rate, and the number of iterations

```
Please input the size of population: 20
Please input the size of Crossover Rate: 0.8
Please input the size of Mutation Rate: 0.3
Please input the mutation selection rate: 0.4
Please input number of iteration: 1000
```

Fig. 2. Input for JSSP using NSGA

3. Result

In the end, it will output the best solution found in all iterations. Since this is a multi-objective problem, there may be multiple sets of solutions. The setting here is to output the same number of solutions as the population size

best obj = [[9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207], [9953.0, 1207]]

elapsed time:33.4816734790802

Fig. 3. Gantt chart for JSSP using NSGA

4. Conclusions

For multi-objective scheduling problems, a nondominated sorting genetic algorithm gives the best solution. The objectives are to minimize the total completion time that means minimize the make-span and total weighted early time and delay time (total weighted earliness and tardiness). These two objectives are conflicting objectives.

References

[1] L. Xu, Z. Y. Xia, and H. Ming, "Study on improving multi-objective flexible job shop scheduling based on Memetic algorithm in the NSGA-II framework," *Proc. 2016 2nd Int. Conf. Cloud Comput. Internet Things, CCIOT 2016*, pp. 1–7, 2017, doi: 10.1109/CCIOT.2016.7868293.

[2] T. Yang, J. Li, and H. Chen, "Multi-objective flexible job-shop scheduling problems with limited resource constraints using nondominated sorting genetic algorithm II," *2018 IEEE Int. Conf. Inf. Autom. ICIA 2018*, no. August, pp. 1127–1131, 2018, doi: 10.1109/ICInfA.2018.8812599.

[3] B. Khan, S. Hanoun, M. Johnstone, C. P. Lim, D. Creighton, and S. Nahavandi, "Multi-objective job shop scheduling using i-NSGA-III," *12th Annu. IEEE Int. Syst. Conf. SysCon 2018 - Proc.*, pp. 1–5, 2018, doi: 10.1109/SYSCON.2018.8369558.

Readiness Factors Influencing the Internet of Things (IoT) in Higher Learning Institutions (HLIs) for E-Learning

Javed Ali[1], Muhammad Shafie Abd Latiff [2], Hafiz Muhammad Faisal Shehzad[3] and Syed Hamid Hussain Madni*[4]
[1]College of Computing Informatics, Saudi Electronic University, Madinah Munawarah, Kingdom of Saudi Arabia.
[2, 3, 4] School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia.
[e-mail: j.ali@seu.edu.sa, shafie@utm.my, muhammad.faisal@uos.edu.pk and madni4all@yahoo.com]
*Corresponding author: Syed Hamid Hussain Madni

Abstract

The increase in university enrolments has led to poor quality education leading to a need for Internet of Things (IoT) as an E-Learning to be utilized in Higher Learning Institutions (HLIs). This enables the institutions to improve teaching and learning through Electronic Learning (E-Learning) during the various situations like Moment Control Order (MCO) or lockdown due to corona. IoT is well implemented in developed countries, however its implementation is reported low in developing countries. There are limited studies that have investigated readiness to implement IoT in E-Learning however the comprehensive identification of factors affecting IoT implementation for E-learning is not yet explored. Therefore, the aim of this research article is to find and analyze the factors influencing readiness for IoT implementation in HLIs. Thereafter, the ease of use, ease of access, technical support and individual skills are significant factors for HLIs readiness to implement the IoT.

***Keywords*:** E-Learning, Internet of Things (IoT), Readiness, Factors, Higher Learning Institutions (HLIs)

1. Introduction

E-Learning is the facilitation of learning by involving electronic devices. Accordingly, this also facilitates expansion of the scope and material of the curriculum [1]. Moreover, it also opens more enrolment spaces for higher learning institutions, universities and colleges. Massive Open Online Courses (MOOCs), virtual learning and flipped classrooms are examples of E-Learning platforms that can provide online learning to students in various situation like distance learning or MCO, etc [2-4].

However, IoT extends online teaching and learning to students and in diversification of processes. It furnishes the E-Learning with intelligence and object integration (things interact to ever object, machines interact with machines) [5]. Briante, et al. [6] also agrees that IoT can make people more knowledgeable about things such as being able to see virtual laboratories from a long distance. The outcome as per the researchers is continuous communication between the real world and the digital world, provisioning the existence of physical objects into the digital world. Moreover,

it can save on costs and help learners to take their classes at any time, within university premises, at home or even in subways [7]. Therefore, IoT is expected to offer solutions that will alter the teaching and learning activity.

IoT as a new actor in learning environments will play a significant role in bringing interactivity, improved learning and understanding between instructors and learners, and virtual and physical objects within the learning environment. Also, there is more focus on smart education and the use of IoT technology to bring improvement to learners in a class [5]. Hence, with an increasing rate to utilize online teaching by higher learning institutions, assessment of readiness in E-Learning is becoming dominant among academics and researchers [8, 9].

Mosa, et al. [10] believe that readiness is a state of being prepared for something about to happen. According to Alshaher [11], this happens at the initial stage of a project for avoiding later risks. Moreover, before acceptance of IoT to universities, consideration should therefore be given to developing the perception and preparation by educators, politicians and the society at large [12]. This is because readiness is more affected by what people believe in, their attitude and intentions.

IoT needs to provide learner support in areas of personalized learning, interactivity, mobility and also accessibility. It can also reduce the education costs and provision for quality education resources than the existing channels. Therefore, the aim of this research article is to identify the readiness factors influencing the IoT in HLIs for E-learning. The rest of research article provides the literature review, methodology adopted, results and conclusions.

2. Related Works

In this section, we review the existing studies that IoT based E-Learning in HLIs. For that purpose Zare, et al. [13] review and classify the exiting research articles that are applied the Multi Criteria Decision Making (MCDM) techniques in E-learning evaluation from the 2001 to 2015. The main objective is to deliver the basic knowledge of MCDM method in E-Learning and also help the further research in the field of E-Learning. Furthermore, Chang [14] presents the comprehensive review and detailed discussion about E-Learning that is used for interactive learning in academia and industry for the training of the staff and employees. Interactive learning integrates the E-Learning and face-to-face learning for ensuring the process of learning that stimulates the learners' interests, progress report, feedback and guidelines to the expected targets.

AjazMoharkan, et al. [15] clarifies the significance and need of IoT, its applications with precise attention on E-Learning. Also, the architecture of IoT, some used and developed technologies, IoT applications in E-Learning and related issues are addressed for further research. In the same way, Moreover, Aldowah, et al. [16] identify the prospective of IoT in HLIs and define to increase the benefit of IoT in E-Learning. Also, challenges and decreasing the risk involved of IoT in E-Learning are addressed. Hence, Zahedi and Dehghan [17] discuss the IoT adoption based E-Learning in several aspects including the benefits, importance, necessity, current challenges, possible solutions and future directions

3. Methodology

A comprehensive review of existing studies is done in the domain of IoT in HLIs for E-Learning to find the readiness factors. The literature from 2016-2020 is included in this research article. The methodology in this research article is based on selection of existing study resources (such as Google Scholar, Science Direct, Scopus, and Web of Science). We use the following search string most suitable for extraction of studies ("E-Learning" AND "Internet of Things" OR "IoT" AND " Higher Learning Institutions" OR "HLIs"). The initial set of studies is examined for removal of duplicate studies, inclusion and exclusion criteria and quality assessment. After that content analysis is performed to categorize the factors and a comprehensive list of E-Learning factors for IoT in HLIs is compiled.

.

Table. 1. E-L-Earning Factors Influencing IoT in HLIs

	References	Almazroi, et al. [18]	Kayali, et al. [19]	Zare, et al. [13]	Chang [14]	Ngampornchai and Adams [20]	Kanwal and Rehman [21]	[Jović, et al. [22]	El-Masri and Tarhini [23]	Naveed, et al. [24]	Hamidi and Chavoshi [25]	Alhabeeb and Rowley [26]	Rahardjo [27]	Naveed, et al. [28]	Muhammad, et al. [29]	Al-araibi, et al. [30]
Individual Factors	Attitude	✓	✓			✓			✓	✓						
	Motivation								✓	✓			✓			
	Computer Competency						✓			✓						
	Anxiety	✓					✓			✓						
	Self-Efficiency		✓			✓	✓			✓						
	Ability			✓								✓				
	Skills									✓			✓			✓
	Behaviour										✓		✓			
	Interaction			✓						✓					✓	
	Feedback									✓						
	Interactive Learning															
Organizational Factors	Course Design			✓		✓				✓				✓		
	User Friendly															✓
	Financial Constraints			✓	✓				✓					✓		
	Infrastructure Readiness									✓						✓
	Financial Readiness								✓	✓						✓
	Training				✓					✓		✓				
	Online Exam/ Test /Quiz					✓										
	Online Monitoring												✓		✓	
Technological Factors	Awareness			✓		✓				✓						
	Multimedia Installment															
	Ease of Access		✓		✓		✓	✓		✓		✓	✓			✓
	Ease of Use	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓		✓		✓
	Technical Support						✓			✓		✓	✓	✓	✓	✓
	Efficient Structure					✓		✓		✓						✓
	System Resources											✓				
	Reliability	✓	✓	✓					✓	✓	✓			✓		
	Time Duration															
	Network Security									✓				✓		✓
Environmental Factors	Support of Faculty				✓							✓	✓	✓		
	Legal and Ethical Issues									✓						
	Class Capacity															
	Progress															
	Tools											✓				
	Meeting Room															
	Internet Speed									✓			✓	✓		

4. Results and Analysis

Several readiness factors are identified that influence on IoT in HLIs for E-Learning as shown in Table 1. We categorized them in four groups according to their nature including the individual, organizational, technological and enviromental factors. Further analysis is discussed below:

4.1 Individual Factors

Individual factors refer to the professional's intention, attitude, motivation and self-efficacy that influence the IoT in HLIs. **Fig. 1**. shows the attitude and self-efficiency are more influenced the IoT in HLIs for E-Learning.

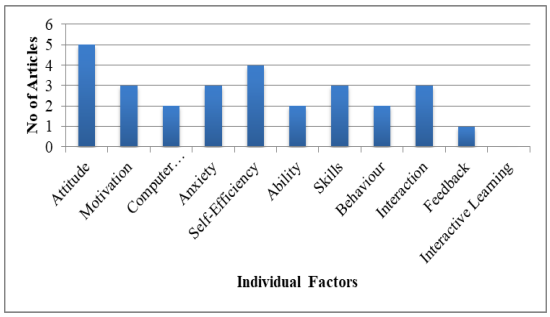


Fig. 1. Individual Factors Influencing IoT in HLIs

4.2 Organizational Factors

Organizational factors relate the inter-organizational processes, practices, and policies that influence the IoT in HLIs. **Fig. 2.** shows the course design and financial constraints are more influenced the IoT in HLIs for E-Learning.

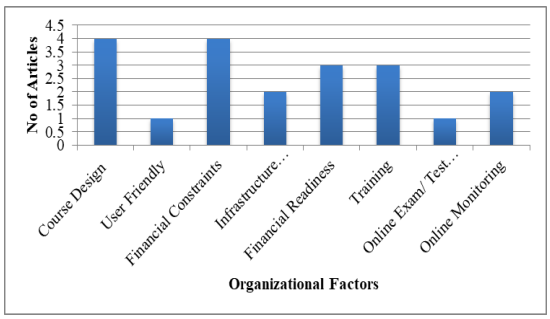


Fig. 2. Organizational Factors Influencing IoT in HLIs

4.3 Technological Factors

Technological factors include the actual use and implementation of technology factors that influence the IoT in HLIs. **Fig. 3**. shows the Ease of use and ease of access are more influenced the IoT in HLIs for E-Learning.

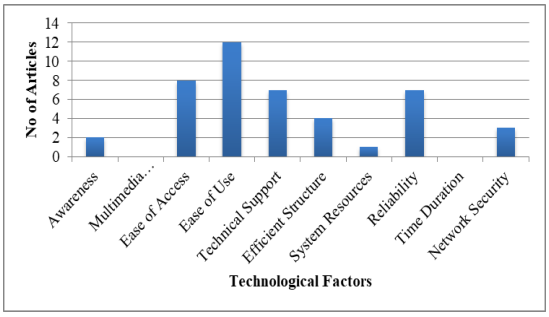


Fig. 3. Technological Factors Influencing IoT in HLIs

4.4 Environmental Factors

Environmental factors are the factors outside of the organization that influence the IoT in HLIs. **Fig. 4**. shows the support of faculty and internet speed are more influenced the IoT in HLIs for E-Learning.

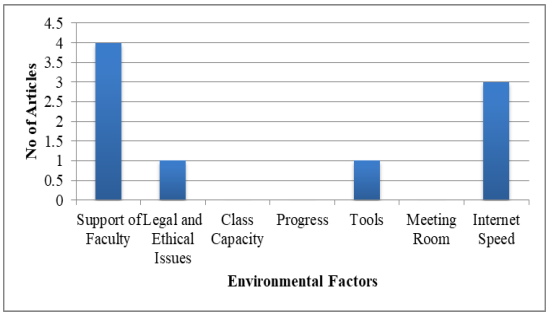


Fig. 4. Environmental Factors Influencing IoT in HLIs

4. Conclusions

Internet of Things can be used as an E-Learning platform to be utilized in Higher Learning Institutions. This enables institutions to improve teaching and learning through E-Learning. This research article aims to find and analyze the factors influencing readiness for IoT implementation in HLIs. The factors are categorized in individual, organization, environment, and technology dimensions. The

finding of the study indicates that technical support, ease of use, ease of access, and individual skills are significant factors for HLIs readiness. Technology developers need to address the complexity of applications and enhance user interaction. Also, there is a need of improving technology skills at universities and higher learning institutions. The quality of service must be improved by technical service providers. The findings are useful for university policy-makers and the government to address the factors and make well-informed decisions regarding IoT implementation for E-learning. Future studies are suggested to develop and validate the IoT readiness model for HLIs.

Acknowledgement

The authors are thankful to Saudi Electronic University (SEU) and Universiti Teknologi Malaysia (UTM) for supporting this research through RUG grant No. ELI-CC120130 and Q.J130000.2451.07G48 .

References

[1] T. D. Baruah, "E-learning as a medium for facilitating learners' support services under open and distance learning: An evaluative study," in *Technology for efficient learner support services in distance education*, ed: Springer, 2018, pp. 93-112.

[2] A. F. Zulfikar, A. Muhidin, W. Suparta, A. Trisetyarso, B. S. Abbas, and C. H. Kang, "The Effectiveness of Online Learning with Facilitation Method," *Procedia Computer Science*, vol. 161, pp. 32-40, 2019.

[3] V. Hamaniuk, S. Semerikov, and Y. Shramko, "ICHTML 2020–How learning technology wins coronavirus," in *SHS Web of Conferences*, 2020, p. 00001.

[4] E. Grabowski, F. Alten, J. Termühlen, P. Heiduschka, V. Brücher, N. Eter, *et al.*, "Analysis of the importance of e-learning in ophthalmology and evaluation of an e-learning app," *Der Ophthalmologe: Zeitschrift der Deutschen Ophthalmologischen Gesellschaft*, 2020.

[5] I. Popchev, D. Orozova, and S. Stoyanov, "IoT and Big Data Analytics in E-Learning," in *2019 Big Data, Knowledge and Control Systems Engineering (BdKCSE)*, 2019, pp. 1-5.

[6] O. Briante, F. Cicirelli, A. Guerrieri, A. Iera, A. Mercuri, G. Ruggeri, *et al.*, "A social and pervasive IoT platform for developing smart environments," in *The Internet of Things for Smart Urban Ecosystems*, ed: Springer, 2019, pp. 1-23.

[7] M. Bayani, K. Leiton, and M. Loaiza, "Internet of things (IoT) advantages on E-learning in the smart cities," *International Journal of Development Research*, vol. 7, pp. 17747-17753, 2017.

[8] A. Ancarani, C. Di Mauro, H. Legenvre, and M. S. Cardella, "Internet of things adoption: a typology of projects," *International Journal of Operations & Production Management*, 2019.

[9] P. Isaias, "Model for the enhancement of learning in higher education through the deployment of emerging technologies," *Journal of Information, Communication and Ethics in Society*, 2018.

[10] A. A. Mosa, M. N. r. b. Mahrin, and R. Ibrrahim, "Technological Aspects of E-Learning Readiness in Higher Education: A Review of the Literature," *Computer and Information Science*, vol. 9, pp. 113-127, 2016.

[11] A. A.-F. Alshaher, "The McKinsey 7S model framework for e-learning system readiness assessment," *International Journal of Advances in Engineering & Technology*, vol. 6, p. 1948, 2013.

[12] F. T. Moreira, A. Magalhães, F. Ramos, and M. Vairinhos, "The power of the internet of things in education: an overview of current status and potential," in *Conference on Smart Learning Ecosystems and Regional Development*, 2017, pp. 51-63.

[13] M. Zare, C. Pahl, H. Rahnama, M. Nilashi, A. Mardani, O. Ibrahim, *et al.*, "Multi-criteria decision making approach in E-learning: A systematic review and classification," *Applied Soft Computing*, vol. 45, pp. 108-128, 2016.

[14] V. Chang, "Review and discussion: E-learning for academia and industry," *International Journal of Information Management*, vol. 36, pp. 476-485, 2016.

[15] Z. AjazMoharkan, T. Choudhury, S. C. Gupta, and G. Raj, "Internet of Things and its applications in E-learning," in *2017 3rd International Conference on*

Computational Intelligence & Communication Technology (CICT), 2017, pp. 1-5.

[16] H. Aldowah, S. U. Rehman, S. Ghazal, and I. N. Umar, "Internet of Things in higher education: a study on future learning," in *Journal of Physics: Conference Series*, 2017, p. 012017.

[17] M. H. Zahedi and Z. Dehghan, "Effective E-learning utilizing Internet of Things," in *2019 13th Iranian and 7th National Conference on e-Learning and e-Teaching (ICeLeT)*, 2019, pp. 1-6.

[18] A. A. Almazroi, H. Shen, K.-K. Teoh, and M. A. Babar, "Cloud for e-learning: Determinants of its adoption by University students in a developing country," pp. 71-78.

[19] M. H. Kayali, N. Safie, and M. Mukhtar, "Literature review of cloud based e-learning adoption by students: State of the art and direction for future work," in *IOP Conference Series: Materials Science and Engineering*, 2016, p. 012087.

[20] A. Ngampornchai and J. Adams, "Students' acceptance and readiness for E-learning in Northeastern Thailand," *International Journal of Educational Technology in Higher Education*, vol. 13, p. 34, 2016.

[21] F. Kanwal and M. Rehman, "Factors affecting e-learning adoption in developing countries–empirical evidence from Pakistan's higher education sector," *IEEE Access*, vol. 5, pp. 10968-10978, 2017.

[22] M. Jović, M. K. Stankovic, and E. Neskovic, "Factors affecting students' attitudes towards e-learning," *Management: Journal of Sustainable Business and Management Solutions in Emerging Economies*, vol. 22, pp. 73-80, 2017.

[23] M. El-Masri and A. Tarhini, "Factors affecting the adoption of e-learning systems in Qatar and USA: Extending the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2)," *Educational Technology Research and Development*, vol. 65, pp. 743-763, 2017.

[24] Q. N. Naveed, A. Muhammad, S. Sanober, M. R. N. Qureshi, and A. Shah, "A mixed method study for investigating critical success factors (CSFs) of e-learning in Saudi Arabian universities," *methods*, vol. 8, 2017.

[25] H. Hamidi and A. Chavoshi, "Analysis of the essential factors for the adoption of mobile learning in higher education: A case study of students of the University of Technology," *Telematics and Informatics*, vol. 35, pp. 1053-1070, 2018.

[26] A. Alhabeeb and J. Rowley, "E-learning critical success factors: Comparing perspectives from academic staff and students," *Computers & Education*, vol. 127, pp. 1-12, 2018.

[27] D. Rahardjo, "E-Learning Readiness and Technology Adoption in Online Tutorial," 2018.

[28] Q. N. Naveed, M. R. N. M. Qureshi, A. Shaikh, A. O. Alsayed, S. Sanober, and K. Mohiuddin, "Evaluating and ranking cloud-based e-learning critical success factors (CSFs) using combinatorial approach," *IEEE Access*, vol. 7, pp. 157145-157157, 2019.

[29] A. Muhammad, A. Shaikh, Q. N. Naveed, and M. R. N. Qureshi, "Factors Affecting Academic Integrity in E-Learning of Saudi Arabian Universities. An Investigation Using Delphi and AHP," *IEEE Access*, vol. 8, pp. 16259-16268, 2020.

[30] A. A. M. Al-araibi, M. Naz'ri bin Mahrin, R. C. M. Yusoff, and S. B. Chuprat, "A model for technological aspect of e-learning readiness in higher education," *Education and Information Technologies*, vol. 24, pp. 1395-1431, 2019.

Integration of Artificial Intelligence in Adaptive Learning Model

Nurul N. Jamal, Dayang N. A. Jawawi, Rohayanti Hassan and M. Irsyad Kamil R
School of Computing, Faculty of Engineering
Universiti Teknologi Malaysia
nnazihah27@graduate.utm.my, dayang@utm.my, rohayanti@utm.my, irsyadkamil96@gmail.com
*Corresponding author: Nurul N. Jamal

Abstract

Besides responsible to help the youth preparing for the unpredictable future, education sectors also responsible to enhance the success of youth in school. Thus, adaptive learning model has become the hot topic among researchers in the recent years. It is suggested in providing good benefits which offered effective approach or methods to improve the youth performance in education. Adaptive learning model provides an environment such personalized learning whereas the learners received learning materials which concerning their preferences. In recommending the learning material, artificial intelligence (AI) technique may be useful in enhancing the development process, including its ability to help in decision-making processes and minimize the uncertainty of sources. However, in which part of adaptive learning model suitable to be integrated with AI so that the data input can be useful for the model. This study focuses on investigating how the recent AI techniques work and applied in adaptive learning model development. In this paper, the use of AI technique is proposed to be integrated at the adaptation model stage to achieve the study objective in recommending learning activities based on the level of knowledge and learning style of the students. The integration of AI technique will make the data input process efficiently and make the adaptive learning model more effective and relevance for teaching and learning process.

***Keywords*:** Adaptive Learning Model, Artificial Intelligence

1. Introduction

Adaptive learning (AL) model has become an attractive subject in the field of education as it is known as systematic approach and technique for improving the performance of students. The AL model can be described as a type of training that uses unique instructional methods or resources to fulfil individuals' specific learning needs [1]. AL needs to support the learners in a meaningful way and to describe success appraisal in the AL software.

In general, researchers designed AL model by creating and developing other sub-models which completed the AL model such as domain model, learner model and adaptation model. For instance, the study done by [2,3,4] present an archictecture which included the three sub-models; (1) domain model, (2) learner model, and (3) adaptation model. All the three sub-models are intertwined to fit various aspects of the instructional process.

Furthermore, in developing AL model, the use of artificial intelligence (AI) is also important to optimize the process on recommending and interprets appropriate materials based on learner preferences. As many human tasks can be carried out by computers today, AI technologies allow a degree of automation that has never been seen within organisations before. AI is a smart

This research was supported by Universiti Teknologi Malaysia – UTMHR Grant Vot No. 08G67

machine designed with human behaviour and reaction [5]. In the recent years, most of researchers [3,6], implement AI technique in classifying and deterimining the learners' learning style and auto recommendation materials.

This study aims to investigate how AI technique be implemented in AL model and in which part of the model is mostly be integrated with. Besides that, this study also investigate what are the recent AI techniques used for the development. This paper is presented as follows: Background of the study is described in Section 2. The findings of the study are presented in Section 3 and the conclusions in Section 4.

2. Background

2.1 Adaptive Learning (AL) Model

A model that being called adaptive in the context of learning if it provides the capacity to include, suggest and intepret relevant materials based on user preferences [3]. As mentioned, generally AL model is designed in the combination of three sub-models which are domain model, learner model and adaptation model as shown in **Fig. 1**.

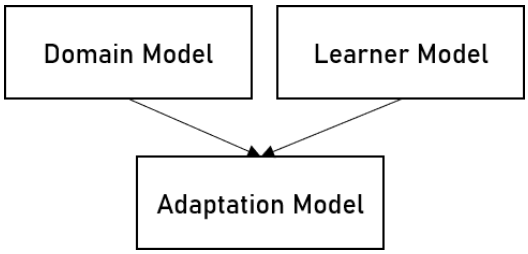


Fig. 1. General AL model

Domain model includes a set of learning resoures, educational materials and attributes, or learning objects [2,7]. Besides organising and store data, this model also promotes different forms of instructional content, not only conventional narrative content but also pedagogical practices. The term of data model, context model, conceptual model or presentation model are also used by researchers to represent this model.

In learning, there must be the learners. In AL model, the information on both static or dynamic data about the learners are vital. It is collected and being stored in learner model [8]. This leaner model work to store student knowledge and awareness of related concepts and preference for learning styles. Besides that, student model term commanly used by researcher to reflect the learner model.

Next, the third sub-model, adaptation model. The intention of the adaptation model is to determine how the expectations of the students towards the presentation of the materials [9]. In order to suggest materials to result good presentation, this model typically consists of a collection of rules that take into account the learner model and domain model. Generally, these three sub-models are vital for developing AL model by most of the researchers. However, it is still depends on what or where the researchers aims to deliver for their adaptive learning projects.

2.2 Artificial Intelligence (AI)

Recommending materials efficiently by considering learner preferences might be challenges. Learner preferences can be form in many facets such as learning style [3], knowledge level [10], experiences [11], learner performance [12] and many more. Thus, AI has been used and integrated to optimize the process on intepretings and recommend appropriate materials which is based on the learners preferences.

In providing learning adaptation, there are different methods of AI that have been used. For instance, the study conducted by Ishak [3] integrated AI techniques at the adaptation model and student model for recommending suitable materials considering students' learning style. Most of existing studies integrated AI in designing AL model. This study integrating AI in adaptive learning model development. Few recent studies are gathered and analysed to see the focus and trends of the integration as depicts in **Table 1**.

3. Artificial Intelligence Technique in Adaptive Learning Model

In the previous section, the background of AL and AI have been described. In this section, the relationship between AL and AI will be

presented. As mentioned, AI is really helpful for automating and optimizing the adaptive learning system especially for recommendation features. The existing studies shows the integration of AI in designing AL model is worked in any part of the sub-model in AL.

Table. 1. Integrated AI in AL model

Ref.	Outcome	AL for	AI in	AI Technique
[3]	Able to deliver learning environment based on learning styles	Learning Environment of materials	Student Model , Adaptation Model	Data Mining Algorithm
[6]	Able to deliver content based on knowledge level and learning style.	Learning materials	Student model	Item Response Theory
[7]	Able to provide learning path based on learning style	Learning path	Domain model	Network-Based Representation
[13]	Able to generate the optimal learning path	Learning path	Adaptation Model	Artifical neural network,
[14]	Able to sequencing the learning material and activities	Learning material and activities	Adaptation Model	Bayesian Network
[15]	Able to determine the sudents learning style and learning model is suggested.	Learning Model	Student Model	Decision Tree, Perceptron
[16]	Able to provide suitable learning material.	Learning Materials	Domain Model	Multilayer Perceptron Artificial Neural Networks

Based on Table 1, it can be seen that, AI can be integrated in any sub-model as illustrated in **Fig. 1**. It is depends on what or where the researchers aims to deliver and focus on their adaptive learning projects. For instance, the study done by [3], AI was integrated in student model and adaptation model as the authors believe the integration of AI at both models able to make the process of development in specifying learning environments of materials more optimize depending on their scope of study.

Most of the studies applied AI at student model and domain model are aimed to analyse and evaluate the student preferences to create student profiles in order to evaluate an overall level of knowledge to be used as the basis for recommend learning pedagogy [3,6,15,16]. The preferences of the students are mostly focus on the their experiences, knowledge of level and their learning style. The learning style can be varies as there are many available learning style model can support the process in completing student model. Based on the study [3,6,15,16] they applied Felder Silverman Learning Style Model.

In addition, AI methods are also often used to encourage the completion of the process so that the learning sources or materials of a courses can be modified to suit each students' preferences. The objective of this study is to investigate does AI technique can optimizely recommend and generate suitable learning activities by considering specific learner preferences if it is integrated at the adaptation model. Based on Table 1, the study done by [3,13,14] also applied AI at the adaptation model. However, the three studies have their own specfic goal and achievement which not suitable in catering the problem of this study. The study done by [13] only tackled the problem of learning object selection based on the mastery level of the learner based on the domain concepts while [14] recommend learning activities based on students experience in completing the activities to investigate the probability of learning activities commanly chosen by the students.

An initial adaptive learning model depicted in Fig 2 is proposed, as an enhancement of the user-intelligent AL model, a study done by Sivakumar et al [17]. It incorporates the three distinct facets of adaptivity. The model consists of three main components: the learner model, the domain model and the adaptation model. As mentioned earlier these components are prevalent to many adaptive learning framework.

In this study, AI will be integrated at the adaptation model to focus on the learning activities selection based on the mastery level and the learning style of the students. The learning style of students will be classified based on Honey and Mamford Learning Style model which complies the Kolb's Learning Cycle principle. AI will be integrated in stage 3 study which is also referred to the study done by

[3,13,14]. The student profile which included the background of knowledge and the learning style of the students, and the learning sources will be the input for the experiment. In AI, there are multiple techniques have been designed and introduced. In this study, classification technique is used for this study based on availability of data. Classification is an AI technique where the data have been categorized into given classes. The main goal of the classification technique is used to identify what are the category of new data of students fall under.

3. Preliminary Result and Discussion

AI technique is proposed to be developed in the adaptation model. As mentioned, classification technique is selected to be used in this study. Thus, a key elements that represent the learning activity repository which included the learners' preferences and conceptual model repository are designed. **Table 2** shows the elements in the learning activity repository. Besides that, a list of rule to categorized the groups based on selected characteristic is set as depict in **Table 3**. The inputs will be processed by applying AI engine (Data mining Algorithm) to result desire output. The proposed output of for this model is the learning activities that suitable with the mastery level and the learning style of the student. **Fig. 3**. Described the proposed pseudo code algorithm for the process in recommending learning activities.

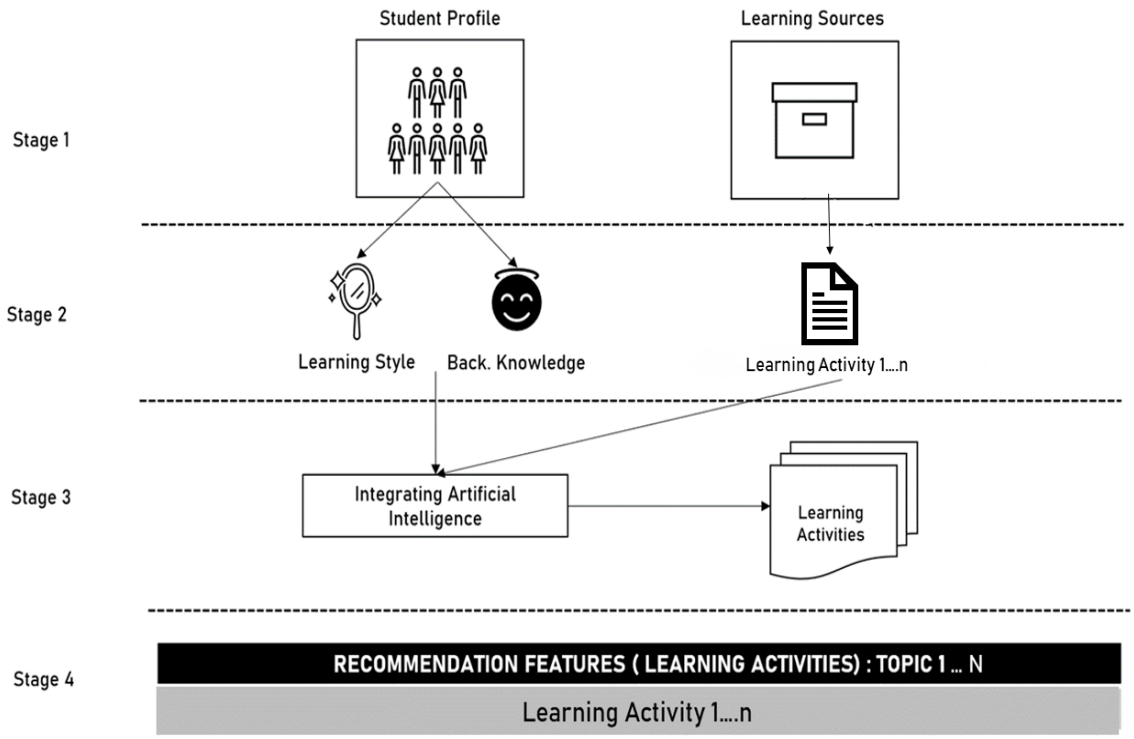


Fig. 2. Proposed Adaptive Learning Model (Initial)

```
function AL-Model [knowledgelevel, learningstyle] return learningactivity

Step 1: Start
Step 2: Read knowledgelevel and learningstyle from student model
Step 3: Read learningactivity from conceptual model where learningactivity related to specific
course domain
Step 4: Repeat until last learningactivity


                    if (student.learningstyle = learningactivity.learningstyle and student.knowledgelevel
                    = learningactivity.knowledgelevel)

                                return learningactivity

                    else

                    if (student.learningstyle = null and student.knowledgelevel = basic/intermediate)

                                if(learningactivity.learningstyle = learning A/b/c/d and
                                learningactivity.knowledgelevel = basic/intermediate)

                                return learningactivity

                                else

                                if (learningacitivity.type = assessment )

                    return learningactivity

Step 5: Stop
```

Fig. 3. AL model Agent Algorithm pseudo code

Table. 2. Elements in the Learning Activity Repository

Dataset	Elements
Dependecies Configuration	**Learning Style (Kolb' Learning Style Cycle ; Honey & Mumford Learning Style Model);** 1: Assimilating (Theorist) 2: Converging (Pragmatist)
	Background Knowledge (mastery level); 1: Basic 2: Intermediate
Conceptal Model repository	**Learning Module;** Module: 1: Introduction to Computational Thinking; 2: Introduction to Robotics; 3: Algorithms and Problem Solving; 4: Programming C with Robotics; 5: Debugging **Learning Outcome:** Learning computational thinking skills through educational robotics problem. **Learning objective:** Able to solve problem given by every modules. **Learning Activity:** A: Lecture B: Brainstorm C: Reflection D:Observation

Table. 3. Set of Rules for Adaptation Model

		Learning Style	
		Activist	Reflectors
Knowledge Level	Basic	Learning Activity: A/B/C/D	
	Intermediate		

4. Conclusions

In this study, the pre-presentation on integration of AI technique in AL model is presented. Futher work need to be done on testing the AL model using a set of student model data while measuring the accuracy of the classification model. The students dataset need to be more precise to make the AL model more effective. Besides that, this AL model will be compared to the other AL model in proving the performance of the proposed method. Other than that, further evaluation is needed to analyse and investigate in which part or stages of AL model development is more suitable to be integrated with AI and the kinds of the AI technique so that the acquired data can be more efficiently used in developing the AL model.

Acknowledgement

We fully acknowledge Universiti Teknologi Malaysia for UTMHR Grant Vot No. 08G67, which have made this research endeavour possible. We would also like to express our sincere gratitude to Software Engineering Research Group members for their continuous support and feedback.

References

[1] H. Peng, S. Ma and J. M. Spector, "Personalized Adaptive Learning: An Emerging Pedagogical Approach Enabled by a Smart Learning Environment," *Lecture Notes in Educational Technology,* pp.

171-176, 2019.

[2] Mahnane, L., Laskri, M. T., & Trigano, P. (2013). A model of Adaptiv e-learning Hypermedia System based on Thinking and Learning Styles. *International Journal of Multimedia and Ubiquitous Engineering*, 339-350.

[3] M. K. Ishak, "Enhancement of Learning Management System with Adaptive Features," UTM, Malaysia, 2016.

[4] T. C. Yang, G. J. Hwang and S. H. Yang, "Development of an adaptive learning system with multiple perspectives based on students' learning styles and cognitive styles," *Educational Technology and S*

[5] M. F. Ahmad and W. R. G. Wan Abdul Ghapar, "The Era of Artificial Intelligence in Malaysian Higher Education: Impact and Challenges in Tangible Mixed-Reality Learning System toward Self Exploration Education (SEE)," in *16th International Learning & Technology Conference 2019*, 2019.

[6] H. Yu and J. Fan, "Design and Implementation of the Framework for Adaptive e-Learning System.," in *Proc. of the Second International Conference on Hybrid Learning and Education*, Macau, China, 2009.

[7] M. Alshammari, R. Anane and R. Hendly, "An E-Learning Investigation into Learning Style Adaptivity," in *The 48th Hawaii International Conference on System Sciences*, USA, 2015.

[8] L. Zaina, "Adaptive Learning in the Educational e-LORS System: An Approach Based on Preference Categories.," *International Journal of Learning Technology,* pp. 341-361, 2011.Molnar and Muntean, 2009

[9] V. Esichaikul, S. Lamnoi and C. Betcher, "Students Modelling in Adaptive E-Learning Systems," *An International Journal (KM&EL),* pp. 342-355, 2011.

[10] N. Aljojo, C. Adams, A. Alkhouli , H. Saifuddin and I. Alsaleh , "In-Depth Analysis of the Felder-Silverman Index of Learning Styles," *American Journal Information Systems,* vol. 3, no. 1, pp. 22-30, 2015.

[11] G. Shelle, D. Earnesty, A. Pilkenton and Erin. Powell, "Adaptive learning: An innovative method for online teaching and learning",*Journal of Extension*,56, 2018.

[12] V. Kasinathan, A. Mustapha and I. Medi, "Adaptive learning system for higher learning," *2017 8th International Conference on Information Technology (ICIT)*, Amman, 2017, pp. 960-970, doi: 10.1109/ICITECH.2017.8079975.

[13] N. Idris, N. Yusof, and P. Saad, "Adaptive course sequencing for personalization of learning path using neural network," *International Journal of Advance Soft Computing and Its Application,* Vol. 1, no. 1, pp. 49-61, 2009.

[14] S. Guiterrez-Santos, J. Mayor-Berzal, C. Fernandez-Panadero, and CR. Kloos, "Authoring of Probabilistic sequencing in Adaptive Hypermedia with Bayesian Networks," *Journal of Universal Computer Science,* Vol. 16, no. 19, pp. 2801-2820, 2010.

[15] R. Bajaj and V. Sharma, "Smart Education with Artificial Intelligence based determination of Learning Style," *Procedia Computer Science,* Vol. 132, pp. 834-843, 2018.

[16] N. Idris, S. Z. M. Hashim, R. Samsudin and N. B. Ahmad, "Intelligent Learning Model based on Significant Weight of Domain Knowledge Concept for Adaptive E-Learning," *International Journal on Advanced Science Engineering Information Technology,* Vol. 7, No. 4, pp. 1486-1491, 2017.

[17] M. S. Sivakumar, M. C. Gombiro and M. S. Venkataraman "A User-Intelligent Adaptive Learning Model for Learning Management System Using Data Mining and Artificial Intelligence,"*International Journal for Innovative Research in Science & Technology,* Vol.1, No. 10, pp. 2349-6010, 2015.

The ways to solve educational and Psychological problems by Based Educational Helping Intervention and Online-Based Emotional Helping Intervention: Focusing on Chinese International Students in Korea

Xiu Jin and Sangwoo Hahm
Department of Business Administration, Honam University, Gwangju, South Korea
Department of Business Administration, Semyung University,Chungbuk, South Korea
[e-mail: soohua1120@gmail.com, thehahm@semyung.ac.kr]
Corresponding author: Sangwoo Hahm

Abstract

International students may face various problems caused by new environments and different cultures when they study abroad in Korea. These problems have a negative impact on the outcome of study abroad. Various studies have been conducted to solve these problems. In order to realize their dream in a global competitive society, the number of foreign students coming to study in Korea is increasing. However, many international students are having difficulties with language problems, cultural differences, and adaptation to life. Such phenomena may degrade the quality of studying abroad and cause failure to study abroad. These negative factors can also have a negative impact on the perception of the local country about the international student and the perception of the international student. Therefore, this study suggests ways to reduce the study-exhaustion of international students studying in Korea. In particular, it is emphasized that educational burnout can be reduced by providing useful information based on online-based helping intervention or e-learning program. International students are divided into three groups - the control group, the educational helping group, and the emotional helping group. The educational helping group provides three weeks of online-based education and stimulation programs. The emotional helping group has three weeks of online conversation, emotional helping, and life information. This research demonstrates how and in what of three groups mental and psychological problems are solved. The purpose of this study is to find out the characteristics of online-based helping intervention and e-learning programs needed for international students. Furthermore, we also find out the way to solve their psychological problems such as educational burnout and stress.

Keywords*:* E-learning program, Online-based helping, Educational burnout, Study abroad,

International student,

1. Introduction

Today, COVID-19 is currently affecting our lives [1] and it also can be seen as a major worldwide threat [2]. Most institutions offer online classes [3]. It is perdict that such situation may lead to a a variety of difficulties to international students who study abroad. For example, online classes can show differences with face-to-face classes. It is believed that international students are worried about their studies. In addition, Cultural differences have a negative impact on international students as well as language and eating habits, as well as psychiatric and psychological problems. Currently, the number of international students studying in Korea is increasing. However, many international students are having difficulties with academic problems, cultural differences, and adaptation to life. Such phenomena may degrade the quality of studying abroad and cause failure to study abroad. These problems have a negative impact on the outcome of study abroad. International students are suffering from various problems. These negative factors can also have a negative impact on the perception of the local country about the international student and the perception of the international student.

In this regard, it can be seen that it is urgent to solve the problems among international students. Various studies have been conducted to solve these problems. This study focuses on solving psychological and educational problems among international students who study in Korea and how to overcome these problems. To solve this problems such as psychological stress and burnout, it focuses on online-based helping intervention and e-learning programs to identify their roles and to verify them through empirical analysis. Hence, this paper explains the strategy and effects of specific online-based helping intervention and e-learning programs.

2. Literature review

Psychological stress is defined as "a reaction to the environment in which there is (a) the threat of a net loss of resources, (b) the net loss of resources, or (c) a lack of resource gain following the investment of resources. Both perceived and actual loss or lack of gain are envisaged as sufficient for producing stress" [4, p.516].

Burnout is a psychological syndrome that involves an extended response to chronic emotional and interpersonal stressors in the workplace. It includes the constant strain caused by a discordance, or unfit relationship between the job and the employee.

It is defined by exhaustion, cynicism, and inefficacy. Exhaustion is the most critical part of burnout, and it is the most evident manifestation. Exhaustion is defined as "a consequence of intensive physical, affective and cognitive strain, that is, as a long-term consequence of prolonged exposure to certain job demands" [5, p. 210]. Cynicism is the development of a negative attitude. This negative reaction is related to the experience of emotional exhaustion. Inefficacy is the tendency to appraise oneself negatively. A person feels unhappy with him/herself and is dissatisfied with his/her own achievement on the job and cannot trust his/her own competence [5, 6, 7, 8, 9, 10]. Through specific education or intervention, the perception of stress and burnout can change, and individuals can become healthier. At this time, education has proven its effectiveness through proven online education methods [11, 12, 13, 14].

3. Method

In this research, the questionnaire is conducted for Chinese students who study in Korea. International students are divided into three groups - the control group, online-based education helping intervention programs group and e-learning programs group.

First, it measures the degree of educational burnout related to the study among Chinese students. They will be provided online-based education helping intervention programs and e-learning programs and they set the period to six weeks. After six weeks, the questionnaire is conducted for Chinese students. Thus, this research will identify the role of online-based education helping intervention programs and e-learning programs by comparing changes in their euducational burnout and stress.

The online-based education helping group

provides six weeks of online-based education trainning. The e-learning program group has six weeks of online chatting, emotional helping intervention, and life helping intervention. For the students who are in control group, there is no intervention for them. In relation to the experiment, firstly, this study measures the level of eucational burnout and stress among the three groups. And then, students who are in online-based education helping group will be provided online-based education and e-learning programs during six weeks. The last, after six weeks, this research will check their degree of educational burnout and stress and how it will be decreased with two online-based programs.

4. Research model

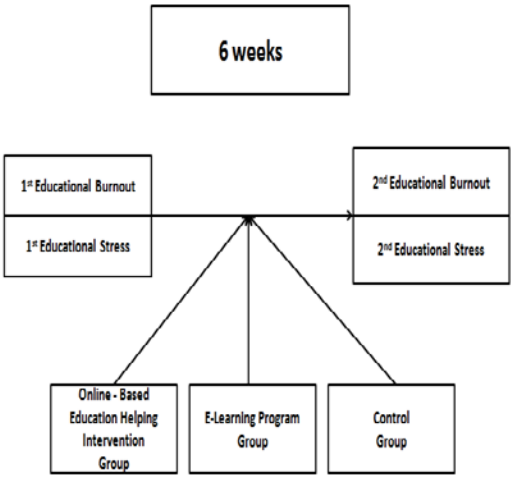


Fig. 1. Research model

5. Conclusion

International students experience various stresses such as academic difficulties, adaptation of culture, difficulty of verbal communication, homesickness, and human relations. The various stresses experienced by international students are linked to their educational burnout. Educational burnout makes people tired and has a variety of negative influences. The performance of international students who are experiencing educational burnout in their study life will be degraded. The performance that educational burnout degrades can include life satisfaction as well as academic performance. The problems of international students who are not satisfied with their lives, or students whose job performance is hindered by educational burnout are not just personal problems. Universitys and the community should help to reduce the educational burnout of international students.

This study aims to find ways to reduce educational burnout and stress by providing help to international students. In particular, online-based helping intervention and e-learning program will slow down educational burnout and stress in each way. First, online-based emotional helping intervention focuses on the emotional difficulties of international students. This approach helps solve the emotional and emotional problems that international students face in their lives. Therefore, we should seek various ways to empathize with the situation of international students and to provide emotional and emotional help. If international students solve emotional problems, they will be better adapted to their study life and educational burnout will be reduced. Second, e-learning program can be a strategy to reduce educational burnout. International students may experience difficulties due to different language and culture teaching methods. Hence, we need the power to help them solve their linguistic problems, to give them opportunities to study with the right colleagues, and to help them understand their major. These strategies will play a decisive role in reducing educational burnout for international students. We will be able to reduce educational burnout for international students through online-based emotional helping intervention and e-learning program. In addition, these efforts will have a positive impact on international students adapting to their study life and achieving high academic performance. These results will help improve the organizational performance of international students as well as international students.

References

[1] A. A. Fernandez, and G. P. Shaw, “Academic leadership in a time of crisis: The Coronavirus and COVID-19,” *Journal of Leadership Studies*, Vol 14, No. 1, pp. 39-45, 2020.

[2] H. Tandon, P. Ranjan, T. Chakraborty, and V. Suhag, "Coronavirus (COVID-19): ARIMA based time-series analysis to forecast near future," arXiv preprint arXiv:2004.07859, 2020.

[3] W. Bao, "COVID - 19 and online teaching in higher education: A case study of Peking University," *Human Behavior and Emerging Technologies*, Vol 2, No. 2, pp. 113-115, 2020.

[4] S. E. Hobfoll, "Conservation of resources: A new attempt at conceptualizing stress," *American Psychologist*, Vol 44, No. 3, pp. 513-524, 1989.

[5] E. Demerouti, K. Mostert, and A. B. Bakker, "Burnout and work engagement: a thorough investigation of the independency of both constructs," *Journal of Occupational Health Psychology*, Vol 15, No. 3, pp. 209-222, 2010.

[6] H. J. Freudenberger, "Staff burn-out," *Journal of Social Issues*, Vol 30, No. 10, pp. 159-165, 1974.

[7] C. Maslach, and E. J. Susan, "The measurement of experienced burnout," *Journal of Organizational Behavior*, Vol 2, No. 2, pp. 99-113, 1981.

[8] C. Maslach, W. B. Schaufeli, and P. L. Michael, "Educational burnout," *Annual Review of Psychology*, Vol 52, No. 1, pp. 397-422, 2001.

[9] W. B. Schaufeli, and P. L. Michael, "Maslach burnout inventory-general survey," *The Maslach Burnout Inventory-test Manual*, Vol 1, pp. 19-26, 1996.

[10] C. Maslach, "Educational burnout new directions in research and intervention," *Current Directions in Psychological Science*, Vol 12, No. 5, pp. 189-192, 2003.

[11] A. J. Crum, S. Peter, and A. Shawn, "Rethinking stress: The role of mindsets in determining the stress response," *Journal of Personality and Social Psychology*, Vol 104, No. 4, pp. 716-733, 2013.

[12] K. McGonigal, "The upside of stress: Why stress is good for you, and how to get good at it," NY: Penguin, 2015.

[13] H. Park, and S. Hahm, "Changes of Stress Mindset and EEG through E-Healthcare based Education," *IEEE Access*, Vol 7, pp. 20163-20171, 2019.

[14] S. Hahm, "Effect of general/job stress mindset to decrease the influence of job stress to burnout," *Information*, Vol 19, No.11B, pp. 5365-5372, 2016.

Secure Password Awareness Educational Gamification System

Anis Erdina Amirullah[1], Cik Feresa Mohd Foozy[1], Nordiana Rahim[1], Adila Firdaus Arbain[2], Palaniappan Shamala[3], Naqliyah Zainuddin[4]
[1]Faculty Computer Science and Information Technology, University Tun Hussein Onn Malaysia, Johor, Malaysia.
[2]School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Malaysia
[3]Faculty Science Computer and Mathematics, Universiti Teknologi MARA (UiTM), Segamat, Johor. Malaysia
[4]CyberSecurity Malaysia, Cyberjaya, Malaysia
*Corresponding author: feresa@uthm.edu.my

Abstract

Gamification is a process of adding games or game like elements to some task to encourage participation. Password is one of the most important safety authentications. Attacks against password is becoming bigger among Internet user. In this paper, the method used is Game Development Life Cycle (GDLC) model. This model is suitable for this project because this model has steps and phases to develop a game. A gamification tool is created to analyze the secure password awareness among university students. After the students have tested the game, they will gain knowledge more about the secure password awareness. Hence, they become aware and more cautious with the dangers and threats against password. The results show that the proposed project is efficient to build the awareness about the secure password among the university students.

***Keywords*:** Attack, Password, Security, Tool

1. Introduction

Gamification is easy to use and one of the learning platforms. Gamification can additionally be used for the purpose of enhancing user engagement and instruction. In gamification, there are regulations to comply with when to gamify a project, lesson, or idea. For instance, rewards can be delivered via leaderboards, badges, and loyalty packages. It can motivate the student to have enjoyable and perform a studying activity. Gamification is not simply about taking part in games, it is also about making sure the student is encouraged to complete the tasks. The student desires the feeling of finished and success of striving towards a challenge.

The password is one of the most necessary matters when we discuss the matter of security. It is genuine that no one can attain a one hundred percent protection level in any area of life specifically now not in information technology. In our digital age, data security is very vital in common and in user authentication. A password is a convenient and handy approach of authentication for users entering a computer system. The system truly requires the user to present something they know as a proof that they are clearly whom they declare to be. This is without difficulty implemented, but at the identical time, the password approach is subject to a number of security threats. If all of these assaults are successful, the account does not

secure anymore. Therefore, it is very vital for the user to create an appropriate and robust password yet effortless to remember. This game-based learning project is to enlarge awareness amongst student about secure password awareness. This kind of awareness will shortly unfold via gamification system as learning to get plenty of attention from the student. It is more advantageous in contrast to campaigns and quizzes. As a result, the student will get encouraged to execute the mission and are trying to find for an answer.

2. Literature Review

2.1 Password

In information security, the password is one of the quintessential and essential aspects. However, human beings occasionally do extra harm than proper with this topic. An unauthorized access has come from a poorly chosen password. It is used to protect user accounts and data.

The more character is using the strong the password is considered. It is additionally necessary for the user to paraphrase their password with the mix of a couple of words and characters. However, it is vital to make the password convenient to remember yet nonetheless be a strong password.

According to [1], the following characteristics are considered as bad or vulnerable password:

1. Contain eight or less characters.
2. Include personal information such as name, birthdate, addresses and telephone numbers.
3. Contain pattern character or numbers such as “abcdefg” or “12345678”.

In order for the user to allow maintaining the multiple passwords, the user can use “password manager” software. Password manager allows the organization to undertake extra tightly closed data-access policies [2](Solaris, Hat, Linux, & Manager, n.d.).

2.2 Password Tool

There are educational games that already exist. Below are the examples of educational games and quizzes that cover the internet safety issues.

1. Webonauts [3] (Niwa, 2007)

Once children have tried on their desired uniform (complete with helmet, of course) they’ll move between worlds and make decisions about password security, social media sharing, bullying, and more. What’s particularly excellent is that the educational aspect is never forced — it feels like a natural part of the story

2. How Cybersmart Are You[3]

This recreation led by way of a cool animated film girl named Sam that will show questions to the users via text and voice. This game covered matters such as passwords, emails, and online ads’ threats. Later, the user will get remarks of their answer whether proper or wrong.

Six papers have been reviewed such as Webonouts, Cyber Protect games, Password Protector game, Elevator game, Password Rehearsal game and CyberCIEGE.

Webonouts is a web-based games. This game is implement for student to have knowledge about password security and media social sharing and bullying. This game interface is enjoyable and colourful. Elevator game is a puzzle game that has been develop to bring discussion about cyber security and to alert about ‘insecure’ behaviour. Password Rehearsal game is a memory games that asked user to memorize the password that has been created. Next is Cyber Protect game, this game is web-based games that can improved user awareness about password, social media security and phishing protection. Password Protector Games is a mobile-based games that train user to create strong password. Last but no least is a simulation games from CyberCIEGE. This game mentioned several types of attack such as spam emails, worms and password cracking for user to learn about the awareness.

3. Methodology

The methodology that is used to conduct this project was Game Development Life Cycle model. This model is suitable for this project because this model has steps and phases to develop a game. The Game Development Life Cycle model has six phases. The phases are initiation, pre-production, production, testing, beta and release phase.

4.Result and Discussion

4.1 Initiation Phase

The first step to do in creating a game is initiation phase. It is to create a rough concept of game that will be developed. The output of initiation is the game concept. User requirement, software and hardware used in this project were also analyzed.

4.2 Pre-Production Phase

Pre-production is one of the first phases in the production cycle. This phase is the planning phase before the production phase starts. Pre-production includes the creation. As a guideline to imagine the route and feel of the amusement, the storyboard is utilized. It also includes the modification of game design and the production of the game prototype. Game design is the most reproachful of the principal pre-production. Game design centers on characterizing game genre, storyline, game features and characters, challenges, fun factors, and player interactions. After the design of the game has been made, a prototype form is made to evaluate the game design.

4.3 Production Phase

Production is the core process in Game Development Life Cycle (GDLC). Production phase was also the longest phase in GDLC. This phase defined the creation of assets, source codes creation, and the integration of both elements. Production activities identified with the creation and the completeness of formal details. It is including balance the game, add new features, improved overall performance, and fix the bugs. Balancing the game implies changes that are identified with the difficulty of the game hard to fit legitimately. Adobe Illustrator was used to designing the game asset. The C# programming language source code was applied to the game asset.

Fig. 1. Home Interface

Fig. 2. Login Interface

Fig. 3. Score Interface

Fig. 1, **2** and **3** shows the interface of the tool. Unity software is used to create the integration between the source code and the game assets. Next phase will test the complete prototype which is the testing phase.

4.4Testing Phase

Testing in this context means internal testing conducted to test the game usability and playability. The testing method is specific to each prototype stage. Formal Details Testing are conducted using playtest to assess the features functionality and the game difficulty (related to balanced). The method to test functional quality criteria is via features playtesting. To test the internally complete quality criteria, it can be done via playtesting simultaneously with functionality test. When a tester discover bugs, loopholes, or dead-ends during playtesting, the causes and scenarios to reproduce the error needed to be documented and analyzed [4]. To test the balanced quality criteria, playtesting with several different treatments is used to categorize whether a treatment is too difficult, too easy, or just fine. Refinement Testing are related to fun and accessibility quality criteria. In refinement testing, fun is tested via playtest and direct

feedback from fellow developers, whether it is boring, frustrating, and challenging. Accessibility can be tested via observing the tester behavior. If tester found it difficult to play and understand the game, it means that the game is not accessible enough. The output of testing is bug report, change request, and development decision. The result will decide whether it is time to advance to the next phase (Beta) or reiterate the production cycle.

4.6 Beta Phase

Beta is phase to conduct third-party or external tester called beta testing. Beta testing still using the same testing methods as the previous testing method, since the related prototypes in the beta testing are both formal details and refinement. The tester selection method comes in two types: closed beta and open beta. In refinement testing, the testers are given more freedom to enjoy the game, as the goals are more directed to get the feedback (related to fun and accessibility quality criteria) Ramadan & Widyani [4]. The output of beta testing are bug reports and user feedbacks. Beta session is closed mainly due to 2 reasons, either the beta term ended or the number of specified beta tester has already given their test report. From here, it may lead to production cycle again to refine the product or continue to releasing the game if the result is satisfactory.

4.6 Release Phase

It is time when the game build has reach final stage and ready to be released to public. Release involves product launching, project documentation, knowledge sharing, post-mortems, and planning for maintenance and game expansion.

5.Conclusion

In conclusion, Secure Password Awareness Educational Gamification System has been successfully developed by accomplishing all the objectives of the proposed gamification system. Moreover, the gamification awareness system able to deliver the importance of secure password awareness in effective way.

Acknowledgement

The authors would like to thank the Ministry of Education Malaysia for supporting this research under Fundamental Research Grant Scheme Vot No. FRGS/1/2018/ICT04/UTHM/03/4 and partially sponsored by Universiti Tun Hussein Onn Malaysia.

References

[1] Sans Institude. (2014). 6 Related Standards , Policies and Processes 7 Definitions and Terms 8 Revision History, 1–2.

[2] Solaris, S., Hat, R., Linux, E., & Manager, P. (n.d.). [Password Manager].

[3] Niwa, H. (2007). 6 Internet Safety Games To Helps Kids Become Cyber Smart. Development, 134(4), 635–646. Retrieved from https://www.makeuseof.com/tag/6-internet-safety-games-kids-cyber-smart/

[4] Ramadan, R., & Widyani, Y. (2013). Game development life cycle guidelines. 2013 International Conference on Advanced Computer Science and Information Systems, ICACSIS 2013, (September 2013), 95–100. https://doi.org/10.1109/ICACSIS.2013.6761558

2D-Based Interactive Media English Learning For Elementary School Students

Endah Sudarmilah[1,*], Taufiq Aryanto[2] and Adila Firdaus Binti Arbain[3]
[1]Informatics Departement, Faculty of Communication and Informatics, Universitas Muhammadiyah Surakarta, Indonesia
Endah.Sudarmilah@ums.ac.id
[2]Informatics Engineering Education, Universitas Muhammadiyah Surakarta, Surakarta, Indonesia
A710150077@student.ums.ac.id
[3]School of Computing, Faculty of Engineering, UTM, Malaysia
adilafirdaus@utm.my
*Corresponding author: Endah Sudarmilah

Abstract

Communication is supported by language. English is one of the most widely used languages in the world and has been established as an international language. Based on research observations conducted at such state elementary school, learning still uses conventional methods and teachers still use worksheets and textbooks to assist learning process. Students only listen and record what is given by the teacher. This study aims to determine the feasibility of the media and the results of the interactive learning media Blackbox testing. Media developed based on 2D for learning English in which there are materials and games that aim to make learning not monotonous. The method used in this research is Research and Development with the SDLC (Software Development Life Cycle) development model (Analysis, Design, Development, Implementation, Evaluation). Media validation is categorized as feasible with an average value of 84.5%. Blackbox media testing test results run well. The results of the media test on the results of the use of instructional media included in the category are quite feasible.

***Keywords*:** english learning, interactive learning media, 2D-based media

1. Introduction

Education is one area of knowledge where the role of technology in the learning process can improve the quality of students learning. In the current era, technology is developing so rapidly that technology is needed in human life, one of which is for knowledge. Multimedia is a new technology that exists in the world of education, where multimedia can be in the form of text, images, animation, video, and sound [1]. In accordance with their age, the education of elementary school children tends to be able to learn faster with material that has pictures, videos, sounds and has attractive colors[2], [3].

English is an international language used in the world, so the world of education emphasizes learning English. Difficulties that are often experienced by children in learning English are difficulties in pronunciation and because English is not an everyday language used. Theoretical learning of English is not effective for elementary school children [4], [5].

Based on the problems that have been explained, it is necessary to develop interactive media for English subjects to facilitate learning in schools. Learning media is the use of tools in the form of objects to help the process of delivering messages [6].

Moreover, learning English still uses teaching methods which uses manual media, students have difficulty on memorizing the arrangement of letters and writing a sequence of letters to form a word[5], [7].

So, the researcher intends to create an interactive learning media for elementary school children by focusing on English subjects with the theme "Fruit and Vegetables", an interactive media created titled "Let's Play and Study" to facilitate teachers in delivering learning material in the classroom.

2. Research Method

This research was conducted using the Research and Development (R&D) method [8]. The product produced in this research is 2D-based interactive multimedia which is used as a substitute for textbooks in elementary school English learning.

This research was designed with the ADDIE development model (Analysis, Design, Development, Implementation, Evaluation)[5], [9]. The flowchart in the ADDIE method can be seen in **Fig. 1**.

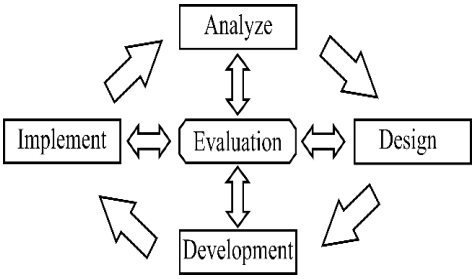


Fig. 1. The Concept of ADDIE

In the analysis phase, the researcher analyzes the problem underlying the making of interactive learning media. The data collection is done by conducting observations and interviews with English language teachers about the lack of media to support learning in the classroom [10]–[12].

At the design stage, design and manufacture of use case, activity diagram and wireframe will be developed [4], [11], [13].

Case Use in 2D-based interactive learning media is made to describe the interaction of users with the system on the media created. In the 2D interactive learning media based on Lets Play and Study, there are game menu, help menu, study menu, settings menu, and credit menu. The diagram can be seen in **Fig. 2**.

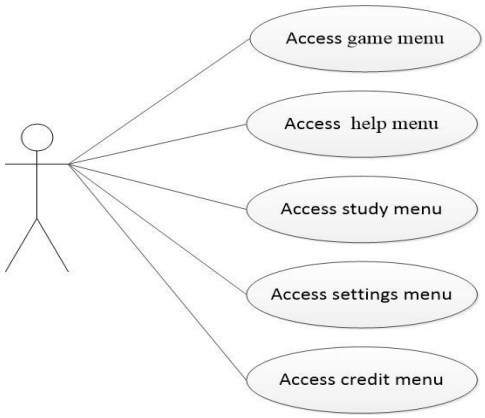


Fig. 2. Use Case Learning Media

Wireframe is the initial design of a product that is made with a picture of the flow of communication, the layout of the image placement and the function of each button before stepping into the manufacturing process. Wireframe in Let's Play and study as **Table 1**.

3. Result and Discussion

Developing of interactive learning media is created using 2D game engine calls by "Let's Play and Study" . On the main page, there are several menus that are displayed include: study menu, game menu, level menu, setup menu, help menu, and credit menu. Which is displayed on **Fig. 3**.

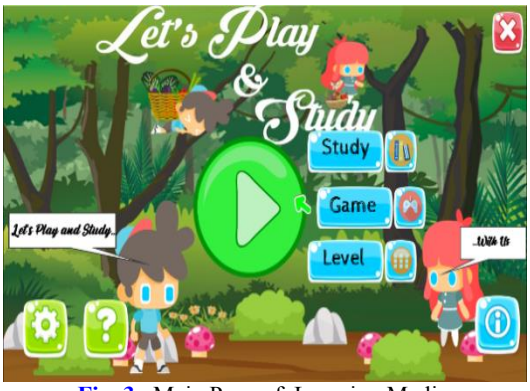
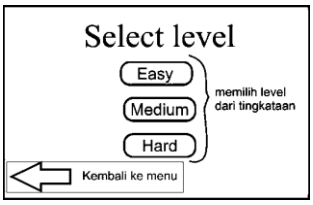


Fig. 3. Main Page of Learning Media

Table. 1. Wireframe in interactive media

No	Design	Detail Information
1	Level Menu 	The level menu has 3 buttons which are used to determine the level in the game. When selecting *easy*, the game will start with an easy level and when it finished, the medium level will open. If the user chooses *medium*, the game will start with a moderate difficulty level and when it finished, the hard level will open. Each level displays the stars obtained, and there is a back button to return to the main menu.
2	Study Menu 	This page shows the English material used. In the application there are voice buttons and English pronunciation. There is also a video button so that the material can be delivered easily, and there is a back button to return to the main menu.
3	Game Menu 	This page is a platform genre game display. The user must find the object that has been determined and head for the finish line by having to pass obstacles and avoid enemies.

The help menu was designed to assist students in operating the media that is used by containing the way to play, the settings menu was conducted to change the sound settings on the media. There are buttons to turn on and turn off the sound on the media. Moreover, the credit menu is designed to display biodata and interactive media creator.

This media was developed with designing the study menu to display English material. By pressing one by one of the learning material, the material contains learning about vegetables and fruit will appear in the info box containing the name of the vegetable or fruit in English. If the image is clicked or pressed, it will sound the vegetables or fruit in English, as seen in Figure 4.

Meanwhile, the game menu was provided to apply the material provided. The game menu displays a game in which 2 characters can be chosen which then collect the names of fruits and vegetables that have been determined. The user collects 3 stars by taking the specified fruit or vegetables and reaches the finish line. In level 2, game turns into a string of letters from fruits and vegetables that have been determined. At level 3, the user matches the fruit and vegetables in the box provided with their names in English as seen in **Fig. 5**.

The level menu was designed to choose based on the difficulty level. There are stars obtained from completing each game as seen in Figure 6.

Software testing focuses on the functional specifications of the software [14]. Blackbox testing in the application Let's Play and study can be seen in **Table 2**.

The media expert testing phase [15], [16] is carried out by filling out the instrument questionnaire provided in order to obtain the data as seen in **Fig. 7**. The questionnaire result data obtained from the experts will be positioned and

used as a presentation of the value as the final result[17], [18]. The formula used is as follows:

$$\text{Percentage of eligibility (\%)} = \frac{\text{(score obtained)}}{\text{(maximum score)}} \times 100\%$$

The percentage of eligibility in **Fig. 1** shows the results of validation by media experts for the developed media. In the aspect of efficiency gets an average value of 87.5%, the aspect of display gets an average value of 81.9%, the technical aspect gets an average value of 83.3%, and the aspect of software gets an average value by 85.4%. The total assessment of aspects of the media expert instrument is 84.5%, so it can be concluded that interactive learning media is quite feasible to use.

Fig. 4. Study Menu Displays

Fig. 5. Game Menu Display

Fig. 6. Level Menu Display

4. Conclusions

The conclusions that can be drawn in the form of applications that are developed are innovations to support learning in the classroom which were originally conventional. The media created makes it easy to learn English by displaying two languages, namely Indonesian and English. The results of the assessment carried out by material experts get an average value of 84.5%. Media created can be run and operated through a PC computer.

Table. 2. The Blackbox testing result of "Let's Play and Study"

No	Item	Case and Testing Result		Result
		Scene	Expected Output	
Main Menu				
1.	Game Menu	Game Button Tap	Switch to the game menu scene	(√) Succeed () Failed
2.	Study Menu	Study button tap	Switch to the study menu scene	(√) Succeed () Failed
3.	Help Menu	Help button tap	Switch to the help menu scene	(√) Succeed () Failed
4.	Level Menu	Level button tap	Switch to the level menu scene	(√) Succeed () Failed
5.	Setting Menu	Setting button tap	Switch to the setting menu scene	(√) Succeed () Failed
6.	Credit Menu	Credit button tap	Switch to the credit menu scene	(√) Succeed () Failed
7.	Exit Menu	Exit button tap	Application exit	(√) Succeed () Failed
Help Menu				
8.	Help Menu	Help button tap	Switch to the help menu scene	(√) Succeed () Failed
9.	Return	Return button tap	Back to main menu	(√) Succeed () Failed
Credit Menu				
10.	Credit Menu	Credit button tap	Switch to the credit menu scene	(√) Succeed () Failed
11.	Return	Return button tap	Back to main menu	(√) Succeed () Failed
Game Menu Game				
12.	Game Menu	Game button tap	Switch to the game menu scene	(√) Succeed () Failed
Study Menu				
13.	Study Menu	Study button tap	Switch to the study menu scene	(√) Succeed () Failed
14.	Sound	Objek button tap	Showing pronounciation	(√) Succeed () Failed
Video Menu				
15	Video Menu	Play tap	Playing learning video	(√) Succeed () Failed
		Pause tap	Learning video pause	(√) Succeed () Failed
		Home tap	Back to material menu	(√) Succeed () Failed
Level Menu				
16.	Level 1	Level 1 button tap	Switch to Level 1, if successfully answer the question, switch to the level 2	(√) Succeed () Failed
17.	Level 2	Level 2 button tap	Switch to Level 2, if successfully answer the question, switch to the level 3	(√) Succeed () Failed
18.	Level 3	Level 3 button tap	Switch to the difficult level, if successfully answer the question, move to the total score	(√) Succeed () Failed
Setting Menu				
19.	Setting Menu	Setting button tap	Switch to the setting scene menu	(√) Succeed () Failed
20.		Unmute button tap	The sound will turn on	(√) Succeed () Failed
21.		Mute button tap	The sound will turn off	(√) Succeed () Failed

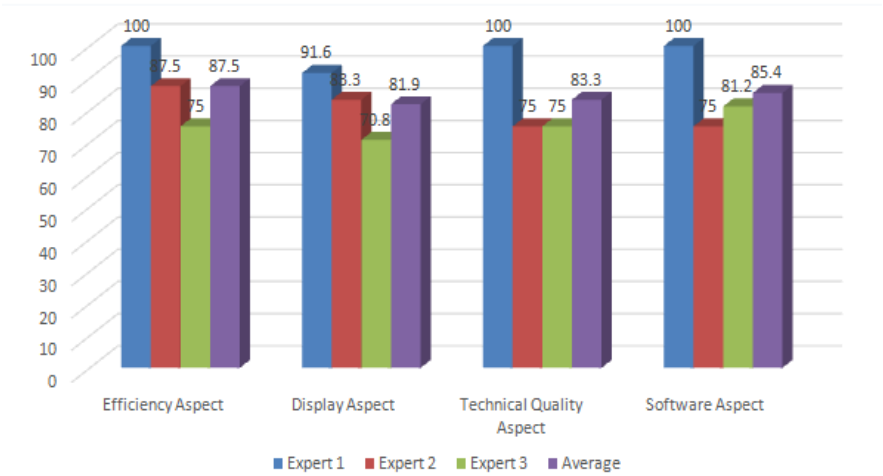


Fig. 7. The result of Media Expert Validation

References

[1] N. Iivari, M. Kinnula, L. Kuure, and T. Molin-Juustila, "Video diary as a means for data gathering with children – Encountering identities in the making," *Int. J. Hum. Comput. Stud.*, vol. 72, no. 5, pp. 507–521, 2014.

[2] O. De Troyer and E. Janssens, "Supporting the requirement analysis phase for the development of serious games for children," *Int. J. Child-Computer Interact.*, vol. 2, no. 2, pp. 76–84, 2014.

[3] E. Sudarmilah, A. Susanto, R. Ferdiana, and N. Ramdhani, "Developing a game for preschoolers: What character, emotion and reward will tend to hack preschoolers?," in *2015 International Conference on Data and Software Engineering (ICoDSE)*, 2015, pp. 89–92.

[4] M. Punt, M. Z. Bjelica, V. Zdravkovic, and N. Teslic, *An integrated environment and development framework for social gaming using mobile devices, digital TV and Internet*, vol. 74, no. 18. 2015.

[5] E. Sudarmilah, Y. Maharani Utami, and I. Yuliana, "3D Interactive Multimedia in English Language Learning," *J. Phys. Conf. Ser.*, vol. 1569, no. 2, p. 022004, Jul. 2020.

[6] A. Ibrahim, F. L. G. Vela, J. L. G. Sánchez, and N. P. Zea, "Educational playability analyzing player experiences in educational video games," *ACHI 2012 - 5th Int. Conf. Adv. Comput. Interact.*, pp. 326–335, 2012.

[7] S. R.-C., P. C., C. M.-M., and H. T.-H., "Effect of Incorporating Project-Based Learning into the Development of an English Short Play for Children," vol. 0. pp. 241–245, 2012.

[8] P. Amarendra Reddy and O. Ramesh, "Security mechanisms leveraged to overcome the effects of big data characteristics," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 2, pp. 312–316, Mar. 2019.

[9] Á. Serrano-Laguna, J. Torrente, P. Moreno-Ger, and B. Fernández-Manjón, "Application of Learning Analytics in educational videogames," *Entertain. Comput.*, vol. 5, no. 4, pp. 313–322, Dec. 2014.

[10] J. O. Grady, *System Requirements Analysis*. San Diego, CA, USA: Elsevier Insight, 2013.

[11] C. Ribeiro, C. Farinha, J. Pereira, and M. Mira da Silva, "Gamifying requirement elicitation: Practical implications and outcomes in improving stakeholders collaboration," *Entertain. Comput.*, vol. 5, no. 4, pp. 335–345, 2014.

[12] G. Jacobs and B. Ip, "Establishing user requirements: incorporating gamer preferences into interactive games design," *Des. Stud.*, vol. 26, no. 3, pp. 243–255, May 2005.

[13] B. Chao, X. Zhu, Q. Li, and A. Huang, "Reliability Management in Software Requirement Analysis," 2006, vol. 2, pp. 1104–1107.

[14] K. Wiegers and J. Beatty, *Software Requirements*. 2013.

[15] E. Sudarmilah, U. Fadlilah, A. Susanto, R. Ferdiana, N. Ramdhani, and P. I. Santosa, "Expert Judgment on Preschoolers ' Cognitive Game Prototype," *Int. J. Pure Appl. Math.*, vol. 118, no. 20, pp. 539–544, 2018.

[16] M. A. Meyer and J. M. Booker, *Eliciting and Analyzing Expert Judgment: A Practical Guide*. SIAM, 2001.

[17] W. Masri and F. A. Zaraket, "Chapter Four - Coverage-Based Software Testing: Beyond Basic Test Requirements," vol. 103, A. M. Memon, Ed. Elsevier, 2016, pp. 79–142.

[18] J. P. Songcuan, A. M. Sison, and R. P. Medina, "Towards usability evaluation of jumbled passsteps," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 4, pp. 1032–1037, Jul. 2019.

The Effect of Job Anxiety by Artificial Intelligence in the 4th Industrial Revolution Era on Job Satisfaction: Focusing on Moderating Role of Job Unclear

Xiu Jin, Chunfan Jin, and Sangwoo Hahm
Department of Business Administration, Honam University
Gwangju, South Korea
Normal College, Yanbian University
Yanji, China
Department of Business Administration, Semyung University
Chungbuk, South Korea
[e-mail: soohua1120@gmail.com, yitswjcb@163.com, thehahm@semyung.ac.kr]
Corresponding author: Sangwoo Hahm

Abstract

According to the advent of the 4th industrial revolution era, our society is rapidly changing. The most of companies are bringing about big changes. As a result of these changes, members are also experiencing various changes, and it is believed that positive changes and negative changes can also accompany them. Artificial intelligence, which is regarded as the core of the Fourth Industrial Revolution, contains both positive and negative effects, and is considered to have both sides. In general, people perceive that artificial intelligence is having a positive effect on humans. Therefore, the positive aspects of artificial intelligence are constantly being emphasized. However, there are also views that the development of artificial intelligence can replace human labor. If members become aware of the possibility of losing their jobs due to artificial intelligence, it may have a huge influence on humans. Therefore, the more constituents who highly perceive their jobs as being replaced, the more they feel the danger of artificial intelligence and eventually perceive anxiety. Job anxiety will generally lead to negative consequences such as career breaks among employees, and job satisfaction is expected to decrease significantly. Therefore, this study emphasizes the negative aspects, unlike research that emphasizes the positive aspects of artificial intelligence according to the changes of the 4th industrial era. The main purpose is to verify the level of recognition of organizational members about negative aspects of artificial intelligence, and to measure anxiety. Furthermore, it presents the they way to solve negative problems in the organization or society to reduce job anxiety.

Keywords*:* Artificial intelligence, The 4th industrial revolution, Job anxiety, Job satisfaction, Job unclear

1. Introduction

As our society in the era of the 4th Industrial Revolution, our society is being influenced positively in various aspects. So people have hope in this positive role. Also, the positive aspects of the 4th industrial revolution are constantly being emphasized. But it doesn't always play a positive role. In some aspects, it

can be seen that side effects can also accompany them. It can be seen as AI as a representative technology of the 4th industrial revolution era. AI is shaping everyday life as fundamental aspects of our society develop rapidly [1]. Therefore, it may have both aspects(positive and negative) will be combined. When it comes to the positive side, AI is making our lives richer. In addition, positive aspects have been emphasized through previous studies [2] [3]. However, in relation to the negative side, it can cause fear under the assumption that machines can think as humans [4]. So, it might think that these negative perceptions are more likely to cause organizational members to think of themselves as useless because of AI. In addition, when job anxiety is high, the level of job satisfaction may decrease due to greater anxiety about the future. According to this, it focuses on the degree to which organizational members feel anxiety due to AI. And the level of job satisfaction accordingly is verified. Furthermore, it is intended to verify the moderating effect of job unclear in the relationship between these two variables. Overall, there are positive aspects of AI, but this study emphasizes that negative aspects are likely to exist as well. Lastly, it suggests that measures are needed to reduce job anxiety among employees.

2. Literature Review

Job anxiety refers to members' disposition to emotional tension caused by anxiety related to one or several factors in their work life, or by common sense of fear and anxiety [5].

Job satisfaction means organizational members have a positive emotional orientation in relation to the organization's employment [6]. Job satisfaction can be seen as an organizational member's positive aspect of their job and overall satisfaction with their job.

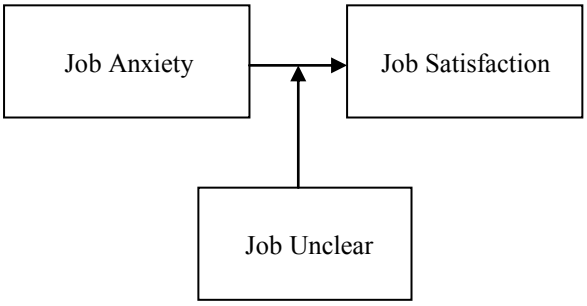
Environmental uncertainty can be seen as frequent and unpredictable changes related to technology development, competitive behavior perceived by managers, and customer preferences [7]. Job uncertainty in the context of changing organizations can be seen as uncertainties about immediate work situations such as promotion opportunities, job changes, and job performance [8, 9].

Hypothesis 1: The perception of job substitution anxiety caused by artificial intelligence will have a negative (-) effect on job satisfaction.

Hypothesis 2: Job uncertainty will have a negative (-) effect on job satisfaction.

Hypothesis 3: Job uncertainty will negatively control the relationship between perception of job substitution anxiety caused by artificial intelligence and job satisfaction.

3. Research Model



4. Results

The perception of job anxiety caused by artificial intelligence lead to low level of job satisfaction.

Job unclear lead to low level of job satisfaction.

Job unclear moderates the relationship between perception of job anxiety caused by artificial intelligence and job satisfaction.

References

[1] C. Cath, S. Wachter, B. Mittelstadt, M. Taddeo and L. Floridi, Artificial Intelligence and the 'Good Society': the US, EU, and UK approach. Science and Engineering Ethics. (2018), Vol.24, No.2, pp.505-528.

[2] R. Lara-Cabrera, C. Cotta and A. J. Fernández-Leiva, A review of computational intelligence in RTS games. In Foundations of Computational Intelligence (FOCI). 2013 IEEE Symposium. (2013), pp.114-121.

[3] L. Rogozea, Towards ethical aspects on artificial intelligence. In Proceedings of the 8th WSEAS international conference on Artificial intelligence, knowledge engineering and data bases. World Scientific and Engineering Academy and Society (WSEAS). (2009), pp.507-512.

[4] A. M. Pereira, M. Gomes and N. C. Fávero, A Inteligência Artificial Utilizada na Recuperação de Informações. Revista Eletrônica de Sistemas de Informação e de Gestão Tecnológica. (2011), Vol.1, No.1, pp.29-44.

[5] S. Verma, Job Anxiety Among Bank Employees. Journal of Psychosocial Research. (2015), Vol.10, No.1, pp.65-71.

[6] J. L. Price, Handbook of organizational measurement. International Journal of Manpower. (1997), Vol.18, No(4/5/6), pp.305-558.

[7] C. Y. Lin and Y. H. Ho, The influences of environmental uncertainty on corporate green behavior: an empirical study with small and medium-size enterprises. Social Behavior and Personality: An International Journal. (2010), Vol.38, No.5, pp.691-696.

[8] J. Buono and J. Bowditch, The human side of mergers and acquisitions. San Francisco: Jossey-Bass. (1989).

[9] S. Jackson, R. Schuler and D. Vredenburgh, Managing stress in turbulent times. In A. Riley & S. Zaccaro (Eds), Occupational stress and organizational effectiveness. New York: Praeger, (1987).

Implementation of NSI Framework on KREONet2: Towards future Internet

Mazahir Hussain[1,2] and Buseung Cho[1,2]
[1] Korea Institute of Science and Technology Information, KREONet Center,
[2] University of Science & Technology,
Daejeon, South Korea
[e-mail: mazahir.hussain@kisti.re.kr, bscho@kisti.re.kr]
*Corresponding author: Buseung Cho

Abstract

Nowadays, big data transfer is critical to enable global research collaboration in data-intensive science like High Energy Physics, Genomics, Climate Change, Astronomy, etc. However, it's too difficult to meet the requirements of the science that needs end-to-end bandwidth-guaranteed network globally using traditional IP routed network. To overcome these challenges, Network Service Interface (NSI) was developed to provide a control framework that dynamically provision end-to-end virtual circuit with bandwidth guaranteed in a global multi-domain network. We developed a testbed to validate OpenNSA, an NSI open source software, and implemented it on KREONet2, the national research and science network of Korea, as one of the network automation solutions. Moreover, to service on-demand end-to-end virtual circuits for data-intensive science globally, we integrated it with research networks of the US such as ESnet, Pacific Wave, and StarLight based on 100Gbps interconnections.

***Keywords*:** data-intensive science, end-to-end virtual circuit, network automation, Network Service Interface

1. Introduction

The collaboration of the research communities is very essential for sharing valuation contents to improve research productivity. These communities need reliable and high-performance computing and networking resources to work on large scale collaborative projects. With the growing nature of networking infrastructure, it is cumbersome to manage the whole network backbone efficiently. Especially, the low-level configuration and troubleshooting errors to normalize the traffic. The challenges are increased for time-constraint, high performance, secure transmission, and on-demand bandwidth for end-to-end multi-domain provisioning of networking resources. Currently, the reservation and provisioning of networking resources are accompanied manually which is time-consuming and complex. Sometimes, misconfigurations lead to network outages. Research communities introduced various frameworks, protocols, and tools for the advancement of current networking architecture.

A framework & protocol is designed by Open Grid Forum with the collaboration of the Research & Education Community and individual to overcome the aforementioned issues such as time-constraint, on-demand bandwidth provisioning of virtual circuits in multi-domain through a secure and high-performance channel known as "Network Service Framework" [1]. Network Service framework documents Network Service Interface (NSI) protocol, Network Service Agent (NSA), Topology Service, and Network Service

for end-to-end network service provisioning. NSA is a software based which communicates with other NSAs using NSI protocol. NSI protocol provides secure and reliable communications between NSAs through well-defined NSI topologies. NSI topology defines the arrangement of the Service Termination points (STPs) where the requests originate or terminate and Service Demarcation Point (SDPs) which connects different network domains. The NSA can be a requester agent (RA) as well as a provider agent (RA) which controls Network Resource Manager (NRMs) in the intra-domain. Multiple NSA can be aggregate by using an Aggregator NSA (AG) to handle requests from child NSAs.

This project intents to investigate the NSI framework on a testbed and deploying our Network Backbone using the OpenNSA [2] Software to collaborate with our international partners for dynamic on-demand circuit provisioning. OpenNSA is an NSI implemented software for dynamic network circuit provisioning. OpenNSA featured with multiple backends such as BROCADE, FORCE10, JUNIPER EX, JUNIPER VPLS, JUNOS EX, JUNOS MX, JUNOS Space, NCS VPN, OESS, PICA8 OVS for intra-domain management.

This paper is organized as follows. The next section discusses related works. Section 3 presents the testbed for the investigation of the NSI framework using OpenNSA software while section 4 demonstrates the current deployment of the OpenNSA/NSI in the KREONet2 (KISTI) network backbone. Finally, Section 5 reviews the challenges, open issues, and future work and concludes the paper in section 6.

2. Related Work

The challenges for research communities increases when it comes to handling big data such as transferring, storing, and processing of big data. However, the processing and storing of big data are handling using robust tools and technologies such as Apache Hadoop, Cloudera, Spark, Cassandra, MongoDB [3-6]. The decreased in the cost of storage and storage technologies is helping to save big data more efficiently [7]. In the collaborations, a highly reliable, efficient, and advanced network is the key to transferring huge amounts of data [8]. It is also required to have a congestion-free and on-demand bandwidth networking resources. Many issues can occur while transferring big data from one site to another site such as server failure, dirty fiber, network congestions, configuration failure, low latency, unexpected errors, etc. A typical organization is bound with the limitation of network infrastructure and resources to handle the transferring of big data. In support of the increasing demand for high bandwidth and reliable networking, different research and education network providers collaborating to overcome this challenge through their backbone [9].

The increasing demand for network resources, Research and Education communities are optimizing, tuning, and enabling their Networking infrastructure for big data transmission and introducing frameworks, protocols, and tools for reliable transmission in multi-domain. Prior works in the advancement of the networking such as Software Defined Network (SDN) [10] to enable programmability and intent-based networking, Data Transfer Nodes (DTNs) and ScienceDMZ [11] for high-performance and advanced technologies in science environment, Globus tools [12] for big data transferring, OSCARS [13] for on-demand secure circuit provisioning in science network, SENSE [14] to build end-to-end deterministic layer2/3 network services, BigData Express [15] for high-performance, predictable and schedulable data transfer, and perfSONAR [16] for monitoring to improve the performance.

3. NSI Framework Testbed at KREONet2

One of the most-efficient ways to test a new framework is by developing a testbed on a real network backbone. The testbed for the evaluation of the NSI framework using OpenNSA software for provisioning of the dynamic circuit in a multi-domain network is shown in **Fig. 1**. In the testbed, we have used two switches Juniper MX80, and MX960 and tagged with 1601-1700 VLAN ids. Each switch MX80 and MX960 is controlled by one OpenNSA as lab1 and lab2 respectively and aggregate both NSA agents using aggregator NSA (AG_NSA).

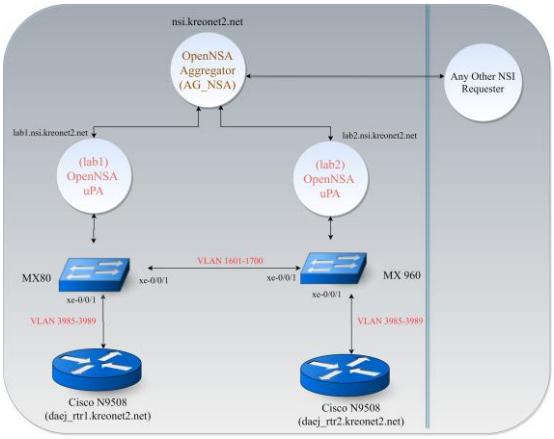


Fig. 1. KREONet2 Testbed for NSI/OpenNSA evaluation

OpenNSA provides eleven primitives to utilize NSI-CS [17] protocols. In **Fig. 2** the command and its respective functions are presented.

```
Usage: onsa [options]
Options:
      --help     Display this help and exit.
      --version
Commands:
    provision             Provision a connection.
    query                 Query a connection (provider summary).
    queryrec              Query a connection (recursive).
    release               Release a connection.
    reserve               Create and commit a reservation.
    reservecommit         Commit a held reservation.
    reserveonly           Create a reservation without comitting it.
    reserveprovision      Create a reservation and provision the connection.
    rprt                  Create a reservation and provision, release and
                          terminate the connection.
    terminate             Terminate a connection.
```

Fig. 2. OpenNSA NSI-CS commands for end-to-end connection provisioning

The result of the 'reserve' query for the reservation of layer 2 connection of 1,000 Mbps bandwidth between lab1 and lab2 domains via 3986 VLAN tags is presented in fig.3. The reservation function needs parameters such as service URL, source STP, destination STP, bandwidth, provider NSA, start and end time to reserve an end-to-end virtual circuit. OpenNSA first checks the availability of the VLAN tags and reservation time before reserving the new connection. On successful reservation, it returns reservation deatails along with connection id after storing in database for provisioning and activating the dataplane. The provisioning function needs parameters such service URL, connection, and provider NSA that activates the dataplane between the STPs. The 'query' command returns the information of the connection that needs parameters such as connection id, service URL and provider NSA [**Fig. 4**].

```
[root@nsi ~]# onsa reserve -u http://nsi.kreonet2.net:9080/NSI/services/CS2 -s
lab1.nsi.kreonet2.net:topology:mx0-0-0#vlan=3986 -d lab2.nsi.kreonet2.net:topology:mx0-
1-0#vlan=3986 -b 1000 -p nsi.kreonet2.net:nsa -a 2020-08-12T06:27:30 -e 2020-08-
12T12:36:00
Connection created and held. Id NS-9919574863 at urn:ogf:network:nsi.kreonet2.net:nsa
Source - Destination: <STP lab1.nsi.kreonet2.net:topology:mx0-0-0?vlan=3986> - <STP
lab2.nsi.kreonet2.net:topology:mx0-1-0?vlan=3986>
Reservation committed at urn:ogf:network:nsi.kreonet2.net:nsa
```

Fig. 3. OpenNSA reservation result between lab1 and lab2 domain connection

```
[root@nsi ~]# onsa query -u http://nsi.kreonet2.net:9080/NSI/services/CS2 -c NS-
9919574863 -p nsi.kreonet2.net:nsa
Query results:

Connection   NS-9919574863 (urn:ogf:network:nsi.kreonet2.net:nsa)
Description  Test Connection
States       ReserveStart, Provisioned, Created
Dataplane    Active : True, Version: 0, Consistent True
Start-End    2020-08-12 06:27:30 - 2020-08-12 12:36:00
Path         lab1.nsi.kreonet2.net:topology:mx0-0-0?vlan=3986 --
 lab2.nsi.kreonet2.net:topology:mx0-1-0?vlan=3986
Bandwidth    1000
Direction    Bidirectional
```

Fig. 4. OpenNSA connection query

Fig. 5 shows the ping results on the successful provisioning of the end-to-end circuit.

```
router2# ping 192.168.104.1
PING 192.168.104.1 (192.168.104.1): 56 data bytes
64 bytes from 192.168.104.1: icmp_seq=0 ttl=254 time=1.21 ms
64 bytes from 192.168.104.1: icmp_seq=1 ttl=254 time=0.794 ms
64 bytes from 192.168.104.1: icmp_seq=2 ttl=254 time=0.707 ms
64 bytes from 192.168.104.1: icmp_seq=3 ttl=254 time=0.883 ms
64 bytes from 192.168.104.1: icmp_seq=4 ttl=254 time=0.739 ms

- 192.168.104.1 ping statistics -
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min/avg/max = 0.707/0.866/1.21 ms
```

Fig. 5. ping test results from lab1 to lab2 after provisioning circuit

4. Implementation of OpenNSA/NSI Framework in KREONet2 backbone

On successfully testing of OpenNSA/NSI on our aforementioned testbed, we move towards the implementation of the NSI framework in our KREONet2 network backbone for dynamic circuit provisioning with our domestic and international partners. The implementation of the NSI framework using OpenNSA is depicted in **Fig. 6**.

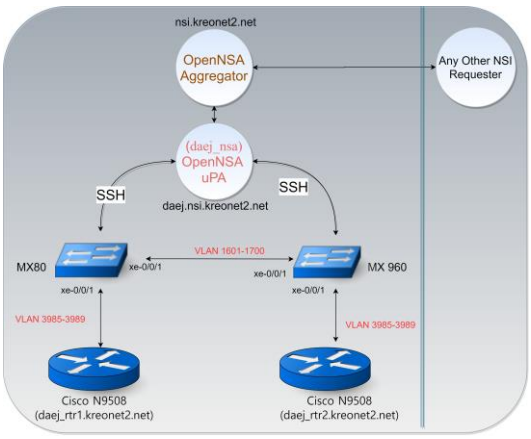


Fig. 6. The model of OpenNSA in KREONet2

In deployment, we have used an Aggregator NSA and one uPA NSA (daej_nsa) to control two switches Juniper MX960 and MX80 enabling 1Gbps and 10Gbps connections. Both switches are connected with our main router CISCO N9508 which connects KREONet2 (Daejeon) to PacificWave and Starlight through our Seattle and Chicago Pops respectively. KREONet2 has 17 domestic and 4 international (Seattle, Chicago, Amsterdam, and Hong Kong) GiGaPops. We have enabled two DTNs through the Cisco N9508 router for transferring data with our international and domestic partners [Fig. 7].

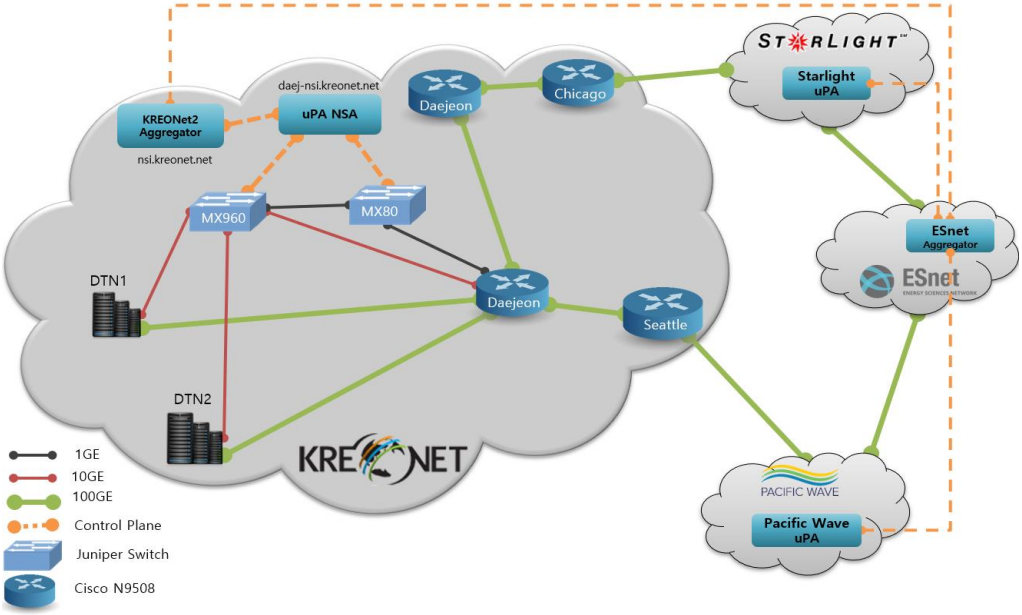


Fig. 7. Enabling NSI framework in production environment with international connectivity

5. Challenges, Open Issues, And Ongoing Work

1. Limitations of current OpenNSA

We identified some of the open issues and challenges after the deployment of the NSI framework in our network backbone. Currently, OpenNSA supports multiple hardware including BROCADE, FORCE10, JUNIPER EX, JUNIPER VPLS, JUNOS EX, JUNOS MX, JUNOS Space, NCS VPN, OESS, PICA8 OVS for intra-domain management. It is challenging to deploy OpenNSA in an environment where the switches and routers are based on CISCO or other hardware that does not support by OpenNSA. We are considering developing for OpenNSA backend for other hardware.

2. Considerations for resource orchestration

The integration of the SENSE system with the NSI framework and bigdata express is one of the future directions, to manage both the networking, computing, and storage resources more efficiently, and to optimize the operation of the networking infrastructure for high-performance data transmission and network orchestration.

6. Conclusions

In this study, we have investigated and deployed the NSI framework using OpenNSA software for the provisioning of the end-to-end virtual circuit dynamically in a multi-domain to overcome the challenges of the on-demand bandwidth, time-constraint, and global multi-domain network. We have enabled highly-optimized and high-performance servers with 100Gbps connections in our network backbone connecting with our domestic research and education institutes and international partners for transferring big data for research productivity.

References

[1] Guy Roberts, T.K., Inder Monga, Jerry Sobieski, John Vollbrecht. *Network Services Framework*. Available from: http://www.gridforum.org/documents/GFD.173.pdf.

[2] *OpenNSA* Available from: https://github.com/NORDUnet/opennsa/.

[3] Khurana, A., "*Bringing Big Data Systems to the Cloud,"* IEEE Cloud Computing, 2014. **1**(3): pp. 72-75.

[4] Gopal, R.C. and A.T. Bharat. "*A generic tool to process mongodb or Cassandra dataset using Hadoop streaming," 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, 2016.

[5] Ghazi, M.R. and D. Gangodkar, "*Hadoop, MapReduce and HDFS: A Developers Perspective,"* Procedia Computer Science, pp. 45-50, 2015

[6] Ahmad, S., P. Singh, and A.K. Sagar. "*A Survey on Big Data Analytics,"* in *2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, 2018.

[7] Siddiqa, A., A. Karim, and A. Gani, "*Big data storage technologies: a survey,"* Frontiers of Information Technology & Electronic Engineering, pp. 1040-1070, 2017.

[8] Settlemyer, B.W., et al. "*A technique for moving large data sets over high-performance long distance networks," 2011 IEEE 27th Symposium on Mass Storage Systems and Technologies (MSST)*, 2011.

[9] *The Global Research Platform*. Available from: http://www.theglobalresearchplatform.net/.

[10] Alam, F., I. Katib, and A. Alzahrani, "*New Networking Era: Software Defined Networking,"* International Journal of Advanced Research in Computer Science and Software Engineering, 2013.

[11] Dart, E., et al., "*The Science DMZ: A network design pattern for data-intensive science,"* Scientific Programming, pp. 173-185. 2014.

[12] Foster, et al. "*Globus Toolkit Version 4: Software for Service-Oriented Systems," Network and Parallel Computing*. 2005.

[13] Guok, C., et al. "*Intra and Interdomain Circuit Provisioning Using the OSCARS Reservation System," 3rd International Conference on Broadband Communications, Networks and Systems*. 2006.

[14] Monga, I., et al., "*Software-Defined Network for End-to-end Networked Science at the Exascale,"* Future Generation Computer Systems, pp. 181-201, 2020.

[15] Lu, Q., et al. "*BigData Express: Toward Schedulable, Predictable, and High-Performance Data Transfer," IEEE/ACM Innovating the Network for Data-Intensive Science (INDIS)*. 2018.

[16] Hanemann, A., et al. "*Perfsonar: A service oriented architecture for multi-domain network monitoring," International conference on service-oriented computing*. 2005.

[17] Roberts, G., et al., *Nsi connection service v2. 0*. OpenGridForum GFD, 2014.

Evolution of Information Exchange Models on Open Dynamic Chatbot Marketplace

KangYoon Lee
Department of Computer Engineering, Gachon University
Seongnam-si, South Korea
[e-mail: keylee@gachon.ac.kr]

Abstract

An App Revolution on the web, on our smartphones, in our cars, on wearable devices, and in our homes has been driving great experiences. Too many apps are everywhere in the marketplace. Even it's hard to remember all how many apps your mobile device has. In the coming era of bots, there are strong requirements that the bot needs to understand which bot or app is appropriate to the user requirements. Bot-based interactive solutions will be placed at the leader of the customer experience in the bot marketplace, just the same as the app used in Apple's store and Google's marketplace. It will evolve into the form of Alexa's skill and Bixby's capsule marketplace. In the case of bots, more than just branching between bot and bot, it can take the form of a multi-turn that invokes the necessary bots to grasp the contents of the information from it and receives the contents and reprocesses them by the caller bot. It takes the form of a universal bot that manages each bot and a bot that handles as a sub-bot for the universal bot. With the bot which has this interface, a new marketplace will be created, and in the future, it will evolve to create the OneBot marketplace, the bots who can communicate between different marketplaces. This study is about the background of the development of OneBot Marketplace and its evolution step. (OneBot marketplace: I named it as the future model of standardized bot marketplace).

***Keywords*:** Chatbot; Bot marketplace; Artificial Intelligence; Cognitive Computing; OneBot marketplace

1. Introduction

App revolution has been making the dramatic changes from the age of PC and Server applications to the current era of apps on mobile devices, now we are transitioning to the era of bots on all instrumented devices which has AI functions as chatbot and interactive dialog service.

Chatbots are developing very quickly and are being applied to various industry on the top of instrumented devices those are PC, Mobiles and all the IoT devices. Many companies such as Amazon, Facebook, IBM, MS and Naver, Kakao are interested in chatbots and have started to support chatbot services on the Cloud PaaS services and in the form of chatbot marketplaces on the top of messenger platforms.

In the past, chatbots used a simple pattern matching method to recognize predefined keywords and output input responses. Recently, chatbots have developed a significant level of AI technology that enables human learning, reasoning, perception, and comprehension. Natural language processing has a very high level of accuracy. Current AI technology is able to understand and respond to the natural language context that humans use, and as the conversation accumulates, the foundation for improving accuracy through self-learning has been laid [1,2]. Chatbot is getting request the development environment and also will have demands of information exchange models between bots.

2. Bot Marketplace

AI Platform companies are driving their own marketplace by development environment with IDE, tools and SDK, those are Alexa Skills, Google Actions and Bixby Capsule.

2.1 Alexa skills

The Amazon's 'Skill' is a store that distributes chatbots through Amazon's 'Alexa' and makes them available to users, and 'Skill' allows users to add features to their AI speaker 'Echo'.

Skills add new capabilities that allow you to create a more personalized experience with your Amazon Echo, Amazon Fire TV, and other Alexa-enabled devices. Skills let you receive flash briefings, order food, request a ride, track your fitness, and more. When you discuss customers setting up Alexa skills, instruct them to "enable" (not download) the skill, and if applicable, direct them to link their accounts in the Alexa app.

The Alexa Skills Store now has more than 100,000 voice apps in September 2019 [3].

2.2 Bixby's capsule

Samsung's Bixby has unveiled its Bixby development IDE and announced plans to deploy Bixby services developed through Bixby capsules and is building the Bixby marketplace.

Bixby Marketplace is a space where users can select and use Bixby Capsule, a service unit available on Bixby-enabled devices. Users can pick and use useful Bixby services in various categories. Bixby Marketplace will provide Galaxy users with more personal and intuitive access to custom AI services. The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

To use the capsule, you can use the service through Bixby by touching the 'Add' button on the Bixby screen. In addition, for some categories, if you set your preferred capsule as the default capsule for a particular command, Bixby will take care of the capsule associated with the command even if the capsule name is not specified. For example, if you set Samsung Music as your preferred music default capsule, say "Play Music" and Bixby will automatically run your search results using Samsung Music Capsule [4].

2.3 Google Actions

Since the launch of the Google Assistant, developer ecosystem has been instrumental in delivering compelling voice experiences. Actions is announced to help that developers build these custom voice apps and services by introducing a suite of new and improved developer tools [5].

3. Evolution of Bot Marketplace

Interactive convergence services such as chatbots are built on the open OneBot marketplace, which supports interoperability between bot platforms. Exchange information between bots will evolve in conjunction with standardized external bot service functions through the connection of Universal-bots and Sub-bots. The development step of market structure is as follows.

3.1. Step I: Individual Platform Ecosystem

According to Bot building, individual marketplace run their own bot application those are Sub-Bot. Each platform has their own IDE and marketplace operation. And there is information exchange between bots. Current Alexa Skill and Bixby capsule marketplace are in the step 1 as shown at **Fig. 1**.

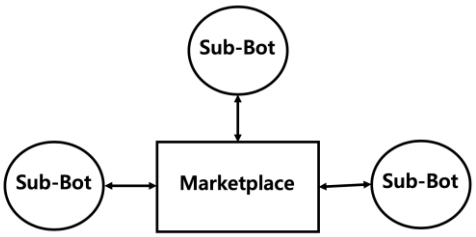


Fig. 1. Step I Market Structure

3.2. Step II : Bot-Bot Communication on Individual Platform

According to bot communication requirements, A bot invokes another bot and exchange information. Caller bot needs to handle sub-bot, which names as Universal bot. Universal bot give the information to the Sub-bot and get results from the Sub-Bot. The operation model shows in the step II structure at Fig2.

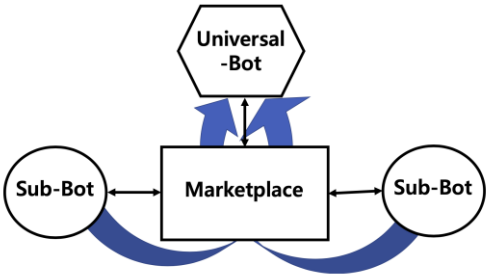


Fig. 2. Step II Market Structure

3.3. Step III : Marketplace with OneBot Platform

Individual Platform start to support Universal and Sub-Bot structure at their own platform which comply OneBot Platform interfaces. However they cannot communicate between bot and bot which exists on different marketplace. The operation model shows in the step III structure at Fig3.

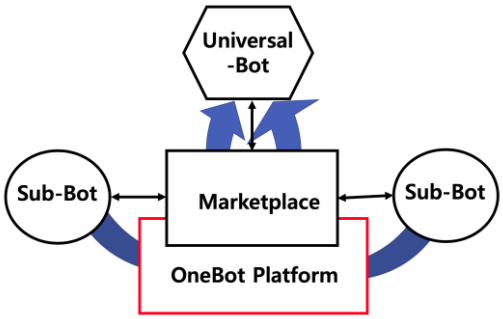


Fig. 3. Step III Market Structure

3.4. Step IV : OneBot Platform Marketplace

As the demand for multiple bots grows, it will support structures that can use bots from other marketplaces. Each marketplace supports OneBot functions to support Sub-Bot by universal-bot's call from other marketplace. The operation model shows in the step IV structure at Fig4.

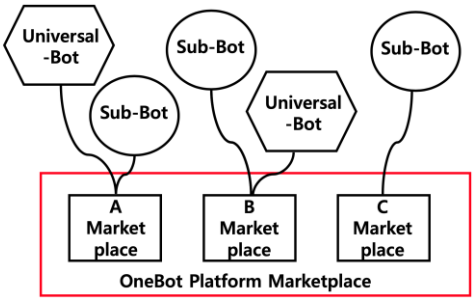


Fig. 4. Step IV Market Structure

In the era of App Marketplace, it is really difficult to invoke other app on the other marketplace. Bot is different from it. Because bot runs on dialog and conversation with information, they can exchange the information. And it will evolve to build OneBot Marketplace

4. Conclusion

The development model of bots and cognitive personal advisors is currently evolving on an interactive basis based on dialog analysis. In addition, interactive visualization techniques according to dialog have been developed to support the derivation of discovery results will evolve into a model that supports the function of the mediator and the debating of the decision-making process [6].

Through the integration of various Sub-Bots Universal-Bot will support user requirements as one bot interface and there will be a new marketplace, aka the open OneBot Marketplace. From now on information structure and protocol for information exchange model between universal bot and sub-bot should be researched and defined. the mediator and the debating of the decision-making process.

References

[1] Abdul-Kader, SA and Woods, JC (2015) 'Survey on Chatbot Design Techniques in Speech Conversation Systems. ' *International Journal of Advanced Computer Science and Applications*, 6 (7)

[2] E. Paikari and A. van der Hoek, "A Framework for Understanding Chatbots and Their Future," *2018 IEEE/ACM 11th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE)*, Gothenburg, 2018, pp. 13-16.
[3] https://developer.amazon.com/en-IN/alexa/alexa-skills-kit.
[4] https://bixby.developer.samsung.com/.
[5] https://developers.google.com/assistant/
[6] KangYoon Lee, "Reference Model and Architecture of Interactive Cognitive Health Advisor based on Evolutional Cyber-physical Systems" *KSII Transactions on Internet*, vol. 13, No 8, 2019

Pose estimation performance analysis of PoseNet model based on edge devices

Min-jun Kim, Young-woo Lee, Hye-min Kwon, Chang-kyo Lee, Jong-han Choi, Min-goo Kang, and Jeongwook Seo
Department of Information and Communication Engineering, Namseoul University
Cheonan, South Korea
Gumi Electronics & Information Technology Research Institute
Seoul, South Korea
Department of IT Contents, Hanshin University
Osan, South Korea
[e-mail: jwseo@nsu.ac.kr]
*Corresponding author: Jeongwook Seo

Abstract

Pose estimation refers to a computer vision technology that detects a person's appearance in images and videos. The resulting Pose estimation model can be applied to the edge devices, and a powerful neural network of the PoseNet model is selected to increase the efficiency and reliability of Pose estimation learning, taking into account various conditions of a general computer environment and other edge devices. There is a need to improve performance on Pose estimation. Therefore, in this paper, we analyze Pose estimation performance for the PoseNet model based on the edge devices. We analyzed the performance of inference times and pose scores for the MobileNet v1 and ResNet50, which are neural networks of PoseNet model implemented in edge devices, and the performance of Pose estimation training according to the device platform version.

***Keywords*:** Edge device, PoseNet, MobileNet v1, ResNet50, Pose estimation

1. Introduction

Pose estimation is the process of predicting the deformation of an object in a user-defined reference pose given an image. This occurs in computer vision or robotics, where the pose or deformation of an object can be used to align computer-aided design models, identify, grasp, or manipulate objects. Implementing image processing using edge computing Implementing image processing software for edge devices By improving the image processing algorithm to suit the edge device environment and improving the image processing algorithm to be mounted on the edge device, edge devices have various restrictions unlike general computer environments[1]. Since there are conditions, it is necessary to improve the algorithm in consideration of this. Therefore, in this paper, we analyze the performance of pose estimation for the PoseNet model based on edge devices.

2. PoseNet model based on Edge device

Fig. 1 explains pose estimation workflow according to PoseNet model. First, the user selects MobileNet v1 and ResNet50 that support PoseNet from the square image, which is the input image, and passes the input image. Through the output model, the reliability score and pose reliability score for 17 keypoints in the input image are returned, and the keypoint heatmap and offset vector are output[2].

Afterwards, without having to go through Pose estimation training one by one, iteratively learns through the pretrained model in the Edge TPU, estimates the pose from the image through the model engine of PoseNet algorithm, and outputs the image from the gstreamer to obtain Pose estimation result[3,4].

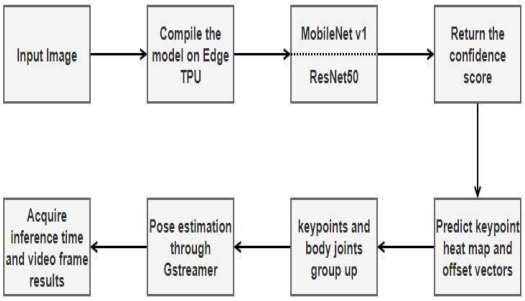
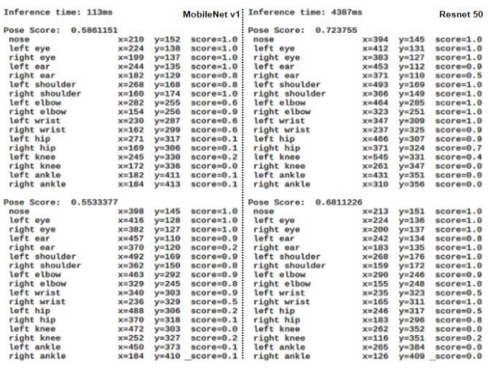


Fig. 1. Pose Estimation Model Workflow for PoseNet Model

3. Manipulation of Virtual Model

Fig. 2 analyzes the performance of pose scores and inference time according to PoseNet model. Reliability scores for each of the 17 keypoints were returned, and one channel per keypoint, x,y, and keypoint location scores were returned. PoseNet model architecture MobileNet v1 and ResNet50 loaded the model of TensorFlow Lite file supported by Edge TPU within the program, and MobileNet v1 showed remarkably high performance in inference time, and the pose score was ResNet50 with a wide layer of each layer. You can see that the performance of is higher.



```
Inference time: 113ms                MobileNet v1 | Inference time: 4387ms                Resnet 50

Pose Score:  0.5861151                            | Pose Score:  0.723755
 nose            x=210  y=152  score=1.0          |  nose            x=394  y=145  score=1.0
 left eye        x=224  y=138  score=1.0          |  left eye        x=412  y=131  score=1.0
 right eye       x=199  y=137  score=1.0          |  right eye       x=383  y=127  score=1.0
 left ear        x=244  y=135  score=1.0          |  left ear        x=453  y=112  score=0.9
 right ear       x=182  y=129  score=0.8          |  right ear       x=371  y=110  score=0.5
 left shoulder   x=268  y=168  score=0.8          |  left shoulder   x=493  y=189  score=1.0
 right shoulder  x=160  y=174  score=1.0          |  right shoulder  x=366  y=149  score=1.0
 left elbow      x=282  y=255  score=0.6          |  left elbow      x=464  y=285  score=1.0
 right elbow     x=154  y=256  score=0.9          |  right elbow     x=323  y=251  score=1.0
 left wrist      x=230  y=287  score=0.6          |  left wrist      x=347  y=309  score=1.0
 right wrist     x=162  y=299  score=0.6          |  right wrist     x=237  y=325  score=0.9
 left hip        x=271  y=317  score=0.1          |  left hip        x=466  y=307  score=0.9
 right hip       x=169  y=306  score=0.1          |  right hip       x=371  y=324  score=0.7
 left knee       x=245  y=330  score=0.2          |  left knee       x=545  y=331  score=0.4
 right knee      x=172  y=336  score=0.0          |  right knee      x=261  y=347  score=0.0
 left ankle      x=182  y=411  score=0.1          |  left ankle      x=431  y=351  score=0.0
 right ankle     x=184  y=413  score=0.1          |  right ankle     x=310  y=356  score=0.0

Pose Score:  0.5533377                            | Pose Score:  0.6811226
 nose            x=398  y=145  score=1.0          |  nose            x=213  y=151  score=1.0
 left eye        x=416  y=128  score=1.0          |  left eye        x=224  y=136  score=1.0
 right eye       x=382  y=127  score=1.0          |  right eye       x=200  y=137  score=1.0
 left ear        x=457  y=110  score=0.9          |  left ear        x=242  y=134  score=0.8
 right ear       x=370  y=120  score=0.2          |  right ear       x=183  y=135  score=1.0
 left shoulder   x=492  y=169  score=0.9          |  left shoulder   x=268  y=176  score=1.0
 right shoulder  x=362  y=150  score=0.8          |  right shoulder  x=159  y=172  score=1.0
 left elbow      x=463  y=292  score=0.9          |  left elbow      x=290  y=246  score=0.9
 right elbow     x=329  y=245  score=0.8          |  right elbow     x=155  y=248  score=1.0
 left wrist      x=340  y=303  score=0.9          |  left wrist      x=235  y=323  score=0.5
 right wrist     x=236  y=329  score=0.5          |  right wrist     x=165  y=311  score=1.0
 left hip        x=488  y=306  score=0.2          |  left hip        x=246  y=317  score=0.5
 right hip       x=370  y=318  score=0.1          |  right hip       x=188  y=296  score=0.8
 left knee       x=472  y=303  score=0.0          |  left knee       x=262  y=352  score=0.0
 right knee      x=252  y=327  score=0.2          |  right knee      x=116  y=351  score=0.2
 left ankle      x=450  y=373  score=0.1          |  left ankle      x=265  y=384  score=0.0
 right ankle     x=184  y=410 _score=0.1          |  right ankle     x=126  y=409 _score=0.0
```

Fig. 2. PoseNet model performance analysis

Fig. 3 analyzes the resolution, inference time, and fps of the pose estimation performance for the device platform.

The PoseNet model used has a lower performance than ResNet50, but the learning layer layer is thin, so we used MobileNet v1, a model suitable for application to edge devices.

The resolution was divided into 640 x 480 and 480 x 360, and the device platform was divided into Raspberry Pi3 and Raspberry Pi4 to check the performance.

The inference time for Pose estimation of the edge devices was 82.0ms based on the Raspberry Pi3 device platform with a resolution of 640 x 480 and the video fps was only 9.27. Raspberry Pi4 device platform with 480 x 360 resolution was the highest at 11.4ms, and video fps was able to measure up to 30.04 fps due to camera performance limitations.

Resolution	Platform	PoseNet	True fps
640 x 480	RaspberryPi3 + EdgeTPU	82.0ms(12.19 fps)	9.27
	RaspberryPi4 + EdgeTPU	20.3ms(49.71f ps)	29.52
480 x 360	RaspberryPi3 + EdgeTPU	43.7ms(22.76f ps)	16.41
	RaspberryPi4 + EdgeTPU	11.4ms(90.23f ps)	30.04

Fig. 2. Analysis of Pose estimation Performance for device platform

4. Conclusions

In this paper, we analyzed the performance of Pose estimation for PoseNet model based on edge devices. Through this, we analyzed the performance of MobileNet v1 and ResNet50, which are PoseNet model architectures, and analyzed the performance of the device platform. Through the experimental results, it was confirmed which of PoseNet models is a suitable model that can consider various constraints on the Edge devices. The algorithm is improved so that Pose estimation model can be mounted on the edge devices, and the edge is different from the general computer environment. We believe this is a model that can take into account various device constraints. Through the analysis of PoseNet model, an image can be saved through the results of Pose estimation in the future, and a system for detecting abnormal behavior can be developed.

References

[1] M. Mukherjee, M. Guo, J. Lloret, Q. Zhang, "Leveraging Intelligent Computation Offloading with Fog/Edge Computing for Tactile Internet: Advantages and Limitations," in IEEE Network, vol. 34, no. 5, pp. 322-329, 2020.

[2] T. Gatt, D. Seychell, A. Dingli, "Detecting human abnormal behaviour through a video generated model," 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA), pp. 264-270, 2019.

[3] T. Gatt, D. Seychell, A. Dingli, "Detecting human abnormal behaviour through a video generated model," 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA), pp. 264-270, 2019.

[4] G. Papandreou, T. Zhu, et al, "Towards Accurate Multi-person Pose Estimation in the Wild" IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

Smart home energy management market segment review

[1*]BadrAlsolami, [2]Muhammad SaiduAliero, [3]YakubuAminu Dodo, [4]Nusa JibrilDanlamil and[5]ErekpitanOmoikhefe Ola-Adisa

[1]Faculty of Islamic Architecture, College of Engineering and Islamic Architecture, Umm-Al-Qura University Saudi Arabia

[2]School of Information Technology, Monash University Malaysia

[3]Department of Architecture, Faculty of Architecture and Engineering, Istanbul Gelisim University, 34310 Istanbul Turkey

[4]Public Building Department, Niger State Ministry of Works and Transport Minna Niger State, Nigeria

[5]Department of Architecture, Faculty of Environmental Sciences, University of Jos, Jos Nigeria

(*corresponding author: bmsolami@uqu.edu.sa)

Abstract

Today, 44% of global energy has been derived from fossil fuel, which currently poses a threat to the inhabitants and well-being of the environment. In a recent investigation of the global demand for energy consumption across various energy consumption sectors, the building sector is one of the primary energy consumers, with a high percentage of energy consumption deemed as unnecessary. This as a result of poor management practice and implementation of strategies to avoid excess energy consumption. This study presents building energy consumption analysis and current smart home energy management products available in the market today to identify current trends and challenges for future improvement. The result reveals a lack of quality attributes such as scalability, interoperability. Lastly, the study described opportunities for future research direction

Keywords: Smart home energy management, smart thermostat, smart home, thermal comfort control, smart energy control, smart home energy management system opportunity

1. Introduction

The global energy crisis has for decades been one of the significant challenges that pose a threat to the global economy and a healthy living environment. A recent report published in 2020 by Energy Outlook [1]shows the increase of 2.9% global energy consumption in 2018, which is considered as the fastest increased rates for decades compared to its usual primary yearlyaverage increase of 1.5% (see **Fig. 1**) and forecasted to increase rapidly in upcoming years. The same report showsthe United States of America was the world's largest energy consumer until 2002 when it was overtaken by the Asian region.

The trend showsa significant turning point closely related to the ongoing increase in population and the number of industries and infrastructures in the Asian region. Another report also shows government efforts to improve the wellbeing and lifestyle of their citizens partly offset by the substantial gain in energy intensity, which leads to higher energy consumption inthe Asian region. A recent report [2] praised China's effort to manage the supply of energy demand to over 1.3 billion people and thousands of industries. Recently, there has been an unusual increase in ongoing deployment of air condition and heating systems, microwave ovens, televisions, hairdryers, and other home appliances.

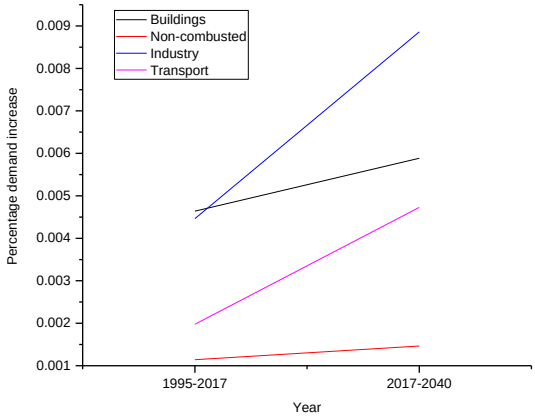


Fig 1. Expected percentage of global demand for energy consumption in various sectors

Currently, the industrial sector is the primary consumer of global energy consumption with increased from 40-90%, buildings (43-56%), transport (18-46%) and the non-combusted accounting for the remainder as shown in **Fig 1.**

The global energy peak consumption analysis from global historical energy consumption data pulled from outlook[1] dated from 1990-2020 in **Fig. 2** indicates Asian region-building energy consumption demand rose by 95% by 2019, consuming a total of 2.26 billion tonnes, overtaking the US with total consumption equivalent 2.17bn tonnes. Similarly, analysis shows a 60% increase in the industrial sector from 1990 to 2019.

Over the years, evidence has shown global warming [1, 3, 4] to be a driving factor that has contributed to a sudden increase inglobal energy demand, particularly in Canada, China, the US, and Russia in efforts to both cool down and warm extreme cold residential environment that has been experienced since the 1950s.

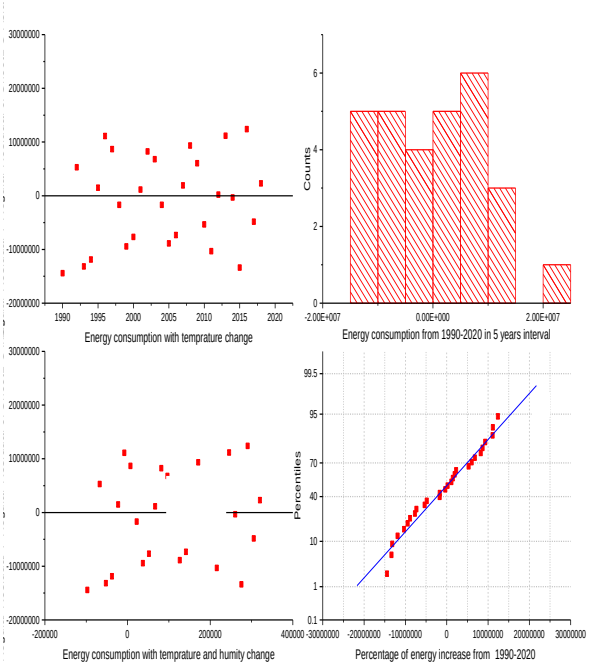


Fig 2. Global energy peak analysis from 1990-2020

SHEMs industries are constantly seeking and demanding for a new recommendation from researchers in both academic and industry to improve the maturity level and smartness of theirtechnologies and their products to keep up and survive with market global market competition [5, 6]and improve the building performance and user experience.

2. Overview of Smart Home Energy Management Market and Challenges

Today the world is in the era of information technology; transforming traditional homes into Smart homes filled with advanced smart networked-based electronic appliances. This innovation allows the resident to control several electrical appliances remotely in achieving a convenient living environment[7].

Current trends[8] show the smart home market rapidly gaining momentum with increased shipments of intelligent home electric appliances globally as a result of its affordability and security. With the current

energy crisis and an increasing population, a smart home has potential to assure quiet, clean, and comfortable environment with available energy supply. Studies show the adaptation of smart homes allows residentsto purchase energy at a low price and to utilize it efficiently, thereby supporting its expansion to the global market. The global market reached $76.6 billion in 2018, and the growth is expected to increase by 28% in 2015. This rapid growth is influenced by crucial factors that include accessibility internet, widespread awareness of the benefit of smart home in ensuring a healthy and comfortable lifestyle environment, improved fitness, simplicity in remote home monitoring and energy efficiency offered. A SHEMS can be divided into independent segments (see **Fig. 3**) serving its purpose [8]as described below:

Smart life entertaining appliances:[8] the key-driven factors that boost the capacity of the smart home market include smart life entertainment appliances. The convenient sound and visual effect, as well as a remote controlling option offered by these appliances, contribute to an increasing number of sales from $170 million to $ $225 billion and is expected to increase by 6.3% from 2019 to 2040. In 2019, Samsung Electronics revealed that all 2018 and 2019 Samsung Smart TV models would enjoy features offered by Apple TV app launched in 2019 in more than 100 countries. With this, Samsung was able to build strong collaboration between Samsung and international investors.

Challenge: The current smart entertainment appliances are not featured to perform proactive, intelligent communication towards the customers. Customers are restricted from getting basic knowledge of both real-time and historical energy data consumption. Integrating proactive features in the upcoming appliances would attract both international and local investors, which help to boost the market share of the smart home in the global market.

Smart lighting system[8]: lighting accounts for large residential energy consumption; however, the current provision of energy-efficient connected lighting controls available in the market reveals the increase in demand for smart lighting system especially for upcoming smart city projects in developing economies. The current market stability has allowed the home resident to plug smart LEDs together with modernized light infrastructure to minimize energy cost and consumption. Today the smart lighting system growth has reached over $20 billion and is estimated to reach $40 billion by 2024. In 2019,the majority of lighting industries acquired technology developers of their Wi-Fi lighting control system. This helps to expand the leadership position and maintain the position in competition with the regional supplier and control the lighting ecosystem price.

*Challenge***:** smart lighting control is one of the fastest-growing segments in the smart home market due to rising demand in automated street lighting, LED occupant lighting systems in developed and Asian Pacific countries. One of the major challenges faced by the smart lighting segment is the compatibility issue. The report showed that in2018, end-users have experienced connectivity issues while connecting their appliance with Google Wi-Fi setup, and it seems the company is aware of the challenge. To survive themarket competition, the smart lighting system has to overcome interoperability and allow connectivity issues across a wide range of technologies.

Integrationof smart lighting with other lighting systems into smart cities such as parking-lot lights, traffic lights, pollution detection sensors, and energy meters is a logical approach to boost the smart lighting share in the market. However, current products available in the market offer only a few devices and accessories to be integrated, which poses a more significant threat to the development of future smart cities. Furthermore, with increased magnitude in voice controlling research, the industry could consider voice control lighting system development solutions together with other innovative lighting applications such as horticulture, speciality, solar, and human-centric lighting. Another opportunity isthe integration of LED wireless-based technology into an automobile. This could serve as a significant advancement for a luxury car to be incorporated with smart lighting.

Smart Access Control: Conventional access cards remain the biggest player in the market. However, today many industries have focused on smart application Bluetooth based technology to offer residents remote access control over secured elevators, doors, and turnstiles[8]. This innovative technology has the potential to revolutionize the digital security market industry eliminating the burden of security access card management and maintenance through the provision of smart access security control via mobile application. Lack of security standard and security architecture requirements to suit current smart devices is one key factor that is slowing the market of smart access control. Even though the smart access control market is still young, recently, the market observed increased demand in smart locks and smart cameras to record and monitor both customer and employee activities. According to the recent forecast on the global crime rate, the estimated market growth is expected to rise from$7.5 billion to $12 billion by 2025. Reliability and affordability of smart access control over traditional proximity and magnetic stripe cards are key driving factors bringing a new investor on board.

*Challenge***:** Although research on IoT in academia and industries is still in its infancy, most of the current access control provides authentication and validation of resident identity without securing end-to-end communication broadcasting resident credentials, among other smart home appliances. Therefore, there is a need for an upgrade on current smart access control to ensure end-to-end security encryption solution to achieve an adequate level of security for the user credentials.

Smart HVAC [8]: is one of the advanced achievements in smart aimed at providing thermal comfort satisfaction to the resident as well as purchasing energy at a low price and avoid unnecessary energy consumption. Today different smart thermos, including occupants' detection, sensors, light, and thermal-based control are available at the global market to ensure stakeholders achieve substantial energy saving and help to obtain real-time information on energy consumption. Recently there is the increased deployment of a smart HVAC platform equipped with sensors that offered remote sensing to regulate the thermostat. This remote sensing application can be configured to monitor and manage room temperature, humidity, and refrigerator. The global market for HVAC control accounted for $14 billion in 2018 and is estimated to reach $27 billion by 2023.

*Challenge***:** Researchers in [4] have shown concern regarding practice employed to regulatesmart HVAC thermostat, whichhas been proven to reduce the life span of smart home energy products. For example, the current method for handling temperature control is mainly based on the classical binary approach, which tends to cause frequent OFF and ON of the appliance when the desire value set is not stable, especially in the afternoon period, which in turn tends to reduce their lifespan.An alternatemethod should be considered,such as a fuzzy control algorithm that would help to regulate temperature efficiently even in a situation wherethe temperature in the room is not stable.

Another challenge is maintaining thermal satisfaction of resident, which is a significant factor that influences the market of smart HVAC system for healthy living and higher productivities at the workplace.Future designs for smart HVAC systems should feature a technique that manages and maintains the thermal comfort satisfaction level of the resident to avoid health challenges.

The Smart Healthcare System:is one of the important IoT technologies that exist today. The technology allows a doctor to treat and care for the patient at home. There is a wide range of applications in the market that enable doctors to examine the health status of the patient at home, such as potential heart attack, drug usage, the blood pressure level of the patient. Other advancements include mechanical applications such as a smart robot, the ambient assistant agent should be a research focal point that would assistin treating viral disease remotely without the risk of being infected. Recently, the University of Tsinghua China developed an automated robot[9] to reduce the spread of the Corona virus among the citizens of China. The robot can determine the potential victim, deliver

food and water to quarantine patients, provide oral treatment. This is one of the major factors that the Chinese government has adopted to reduce the spread of the Corona virus to medical personal.

Today there is increased demand for smart mechanical healthcare assistant agents to monitor and assist patients suffering from diabetes, heart ailments, asthma, and joint pain. The global market for smart healthcare system already hit $6 billion and forecast to reach $30 billion by 2026. The major key players in this sector include Health Care Originals, Apple, Google, and Medical Guardian LLC.

Challenge:The smart healthcare industry is one of the fastest-growing segments in the smart home market today. This is driven by improving ease of access to healthcare application cost. However, with the current increase in the world population, the outbreak of viral disease and an increasing number of older people requires an ambient assistant. There is a need to increase the number of innovative applications to provide autonomous assistant to treat the patient and assist older people.

Current smart healthcare applications require integration of many technologies to connect patient with doctors, and on many occasions, these applications face interoperability challenges. Researchers in academia and industries must focus on developing cross-platform layers that address interoperability issues in smart healthcare applications.

Smart kitchen[8]:uses different sensors designed to provide comfortable and convenient kitchen activities. Theseappliances are well equipped with wireless and Bluetooth based connectivity features to provide communication with other smart house appliances such as tablets and other remote handheld appliances. Today, academia and industries are focusing on smart kitchen integration and other start-up products with several smart kitchen appliances available in the market equipped with sensors for easy operation and to help residents regulate kitchen activities remotely. Smart kitchen appliances are energy efficient compared to conventional kitchen appliances, which is expected to raise their demand in the market in the coming years. The global market growth for the smart kitchen has already reached $2.7 million and expected to reach $8.5 million by 2027. The major participants in the market include LG (US) and Tovala.

Challenge: Whenever we talk about the smart kitchen, everyone will think of a phenomenon where you have kitchen system control by artificial intelligence to prepare a favourite resident meal. Several smart kitchen appliances available at the market today are handy. With today's reliance on technology, it is a big opportunity for smart kitchen industries to consider a more sophisticated smart kitchen that would adapt to a favourite resident meal based on resident historical data for more straightforward preparations. Featuring the next smart kitchen generation with the ability to recognize your favourite meal, order recipe and ingredient and cooked resident favourite food will the game-changer in the smart home market.

Smart Furniture[8]: is the segment of smart home solutions that monitor residents surrounding information to ensure comfort satisfaction and integrated functionality. An increase in individual income and the nation's economy is a key factor influencing individual lifestyle change. The technology mainly monitors resident fitness such as bedtime, nutrition, hygiene, and total burned calories. Additionally, smart furniture has features that provide wireless charging to smartphones and Bluetooth speakers. Recently it was announced the upcoming smart furniture would feature technology that monitors employee work productivities, the status of the closet, and alarm for the messy environment. The global market for smart furniture has reached $174 million and estimated to hit $795 million by 2026. The market players in smart furniture include Smart Living LLC, Ori Systems, and Ikea Group.

Challenge: The current smart Furniture available in the market can monitor the status activity style of the resident within the surrounding which includes standing, seating, sleeping and eating, and this information is being shared among other smart home appliances such as smart home meter

autonomously without any authentication mechanism. This practice also can introduce a privacy challenge that can lead to theft or committing a high-level crime since criminals can manage to gain access to smart furniture and deduce the time the resident is at sleep or not present at home. The industries should consider advanced security mechanisms in future smart furniture to ensure properidentification before the establishment of communication betweensmart home and outsider.

3. Discussion and future research direction

Recent developments[10, 11] of SHEMs products show the increasing shipments smart entertainment, HVAC, lighting, healthcare, kitchen globally. Many of these ongoing deployments have not attained a mature level of smartness, provide meaningful synergy between humans and appliances, primarily with applied AI algorithms. For example, embedding virtual assistant, voice control AI to smart furniture, kitchen, andlighting appliance can provide additional functionality such as making calls, activating, and connecting other devices to enhance collaboration between home appliances and occupants. The adoption of this collaboration mechanism can significantly improve home occupant's comfort, lifestyle experience, and energy-saving potential.

Using advanced AI to establish collaboration between smart furniture and kitchen appliance. Before occupants leave the bed, AI assistance can autonomously communicate with kitchen appliances to begin preparation of occupant's favourite mills by accessing the build-in recipes gallery before. It can also alert the occupants when a recipe is low of stock. Adoption of this AI advancement can minimize an average of 30% of cooking energy consumption[12].

Embedding deep learning in smart home appliances can enhance their simplicity and occupant's experience by learning and adapting to occupant's daily scheduled activities. For example, if the occupants set a gym workout on the calendar, the washing machine should be able to turn ON gym clothes wash ON as the occupant returns home. Improving user experience using the Fuzzy logic system to ensure smart sensors detect the laundry type water level automatically required the clothes are placed in the washing machine to avoid unnecessary energy consumption and water waste. Improving the washing machine controller to schedule cloth washing activities or notify occupants when the cost of energy is low. This practice will avoid operating appliances during peak energy demand at a higher price[13].

Many of the mart furniture products available in the market today[12], such as a smart speaker, TV are implementing AI-based technology example include Google Home or Amazon Echo that use voice recognition software to feature voice control assistant via mobile app or voice commands to perform many tasks like make grocery lists, create a playlist, search the web and even turn ON reminders. These appliances require serious security consideration as they can easily be accessed and control via the internet.

References

[1] Outlook, E., *Global energy statistical yearbook 2019.* Enerdata, 2019.

[2] 2018, C., *How is China's energy footprint changing?* Online Report, 2018.

[3] Part, B.W.E.O., *World Energy Outlook.* 2016.

[4] Serra, J., et al., *Smart HVAC control in IoT: energy consumption minimization with user comfort constraints.* ScientificWorldJournal, 2014. **2014**: p. 161874.

[5] Ullah, F., F. Al-Turjman, and A. Nayyar, *IoT-based green city architecture using secured and sustainable android services.* Environmental Technology & Innovation, 2020. **20**.

[6] Haseeb, K., S. Lee, and G. Jeon, *EBDS: An energy-efficient big data-based secure framework using Internet of Things for green environment.* Environmental Technology & Innovation, 2020.

[7]. Darianian, M. and M.P. Michael, *Smart Home Mobile RFID-Based Internet-of-Things Systems and Services*, in *2008 International Conference on Advanced Computer Theory and Engineering*. 2008. p. 116-120.
[8]. MarketsandMarkets, *Smart home market by product.* Online Report, 2020.
[9]. Gangtie, Z., *Robot to perform examinations on patients infected with the coronavirus.* Online Report, 2020.
[10]. Shin, J., Y. Park, and D. Lee, *Who will be smart home users? An analysis of adoption and diffusion of smart homes.* Technological Forecasting and Social Change, 2018. **134**: p. 246-253.
[11]. Marikyan, D., S. Papagiannidis, and E. Alamanos, *A systematic review of the smart home literature: A user perspective.* Technological Forecasting and Social Change, 2019. **138**: p. 139-154.
[12]. Guo, X., et al., *Review on the Application of Artificial Intelligence in Smart Homes.* Smart Cities, 2019. **2**(3): p. 402-420.
[13]. Mazhar, M.H. and Z. Shafiq, *Characterizing Smart Home IoT Traffic in the Wild*, in *2020 IEEE/ACM Fifth International Conference on Internet-of-Things Design and Implementation (IoTDI)*. 2020. p. 203-215.

Review on smart home energy management system approaches

[1]Yakubu Aminu Dodo, [2]Muhammad Saidu Aliero, [3] Amina Batagarawa, [4]Soumia Mounir and [5]Lawrence Pendo Moveh
[1]Department of Architecture, Faculty of Architecture and Engineering, Istanbul Gelisim University, 34310 Istanbul Turkey.
[2]School of Information Technology, Monash University Malaysia.
[3]Department of Architecture, Faculty of Built Environment Science, Baze University, Abuja-Nigeria
[4]National School of Architecture FEZ, Parc FEZ Shore, FEZ Morocco New university complex, Agadir 80000 Morocco.
[5]Bola Tinubu Health & Diagnostic Center, Lagos State University Teaching Hospital (LASUTH)Lagos State, Nigeria.
(*corresponding author yadodo@gelisim.edu.tr)

Abstract

A smart home energy management system is the advancement of the technological concepts of the Internet of Things (IoT) aimed to reduce excess energy consumption in buildings. Over the years, researchers have proposed different techniques and strategies to predict the possibility of building occupants to avoid unnecessary heating/cooling of vacant space. This study conducts a review analysis on the subject of smart home energy management systems (SHEMS) published between 2010 and 2020. The study presented classification of SHEMS approaches and findings reveals there are more proposal algorithms to schedule energy consumption by appliances during when the cost of energy is too high or demand is at peak to minimize the cost and ensure the sustainability of utility. Further, the analysis reveals research focuses on the implementation of algorithms in a commercial building using an occupant's fixed schedule, which might produce a poor result when implemented in residential buildings. It further reveals camera-based image and video processing strategy is one of the most efficient techniques so far proposed to bridge the gap between HVAC energy-saving and acceptable indoor thermal comfort level. However, it significantly suffers from privacy issues.

Keywords : approach, energy-consumption, HVAC & SHEMS

1. INTRODUCTION

Global warming is being widely deemed to be the biggest threat facing mankind [1, 2]. Despite the Paris Agreement, the Inter-governmental Panel on Climate Change warns that the world is well short of making due progress towards limiting the rise of temperature to under 2℃. A recent IPCC report demands revamping the four big global systems: energy, land use, cities, and industry. The building sector has to play an important role in this respect as, globally, it accounts for over one-third of the energy consumption, greenhouse gas emissions, and around 40% of the consumption of natural resources[3, 4]. Worldwide, the building industry is pursuing sustainable measures to curtail associated energy and environmental burdens. Sustainable energy solutions in the form of energy efficiency and renewable energy are playing a critical role in this respect[4, 5]

The building industry is actively seeking new technological solutions to improve their energy performance. Smart home energy management system (SHEMS) is one of the

advancements of the Internet of Things (IoT) technologies that emerged in the global information industry, equipped with IoT products and ZigBee technology aimed to transform the traditional building into an energy-aware environment to minimize excess energy consumption. The major contributors to excess home energy consumption include Heating, Ventilation, and Air conditioning (HVAC) systems and lighting, which are necessarily required to maintain a healthy living in an indoor environment[6, 7]. HVAC accounts for as much as 80% of the total energy use in a building in harsh climate conditions, while lighting accounts for around 12% of the entire load [8]. A significant proportion of this total energy consumption is avoidable through effective management of these energy requirements, especially when the concerned space is unoccupied.

Reliable and accurate human occupant's detection in an indoor space is one of the influential factors that help to control the use of energy during the period of space vacancy. The research on occupant detection integration to manage HVAC operation has received attention in recent years targeting both commercial and residential buildings to reduce unnecessary energy consumption. A recent study has shown an average of 25% energy saving potential through HVAC equipment in comparison with traditional approaches[9].

2. Classification of SHEMs Approach

Current literature in the smart home energy management system can be classified into predictive and non-predictive approaches presented in **Fig. 1**.

Predictive approach: this is a type of control designed to automate the management of HVAC operation based on occupants' data and indoor meteorological data collected from sensors to estimate the probability of a building being occupied. These occupant data can be static (fixed schedule) or dynamic real-time data collected by installed sensors in space.

Static predictive control: deals with literature that greatly relied on occupant's assumptions and fixed schedule activities to derive a model that estimates the probability of a building been occupied with controlling HVAC operation. This type of control approach can be practical in a space where occupant activities are governed by a fixed schedule policy strictly being followed daily. A good example is commercial buildings such as labs, offices, and business environment. However, this approach might produce a poor result in a space where the occupants do not follow a fixed schedule.

Dynamic predictive control: This class of control heavily relies on real-time occupants data to develop predictive model-based optimization and an expert system such as rule-based or knowledge-based, algorithm. The study shows this class of control have been severally deployed by different researchers to control HVAC indoor ventilation which relatively offers easy tuning and can cope with the multi-variable problem such as occupant's number, sudden change of ambient temperature or humidity. However, its accuracy is less reliable to constantly react and to the change in indoor air quality to maintain satisfactory comfort level, and location of occupants to reduce unnecessary energy consumption.

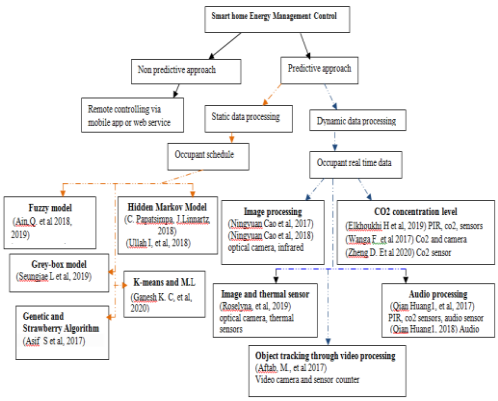


Fig 1. Classification of smart home energy management system approaches

The current SHEMs building research has shown the strong emphasis and momentum on commercial buildings compared to residential

buildings as shown in **Fig. 1**. The strategies used to manage HVAC operation in most of the commercial building relays on occupant's assumptions and schedule activities are usually extracted from the organizational policy. This approach might be unreliable in a situation where the occupants deviate from certain scheduled activities and could not guarantee a satisfactory comfort level or respond to change of indoor air quality due to its static nature.

Another major challenge across a fixed schedule is the lack of ability to satisfy the occupant's thermal comfort as a result of fixed setpoint value temperature, which has been used throughout the day. A recent investigation [10-12] of thermal comfort satisfaction reveals the acceptable level of air quality level essential to maintain a healthy indoor environment. The study investigation reveals only a few of the existing control (see Table 6) attempt to bridge the gap between thermal comfort and energy saving while the majority of the study focuses on energy-saving potential ignoring the thermal comfort.

2.1 Technologies used in SHEMS

The final study selection examined literature that attempts to bridge the gap between thermal comfort and energy-saving potential. The study analyses selected literature [8, 10, 12-19] based on technological and research challenges and a summarised the findings

The study observed motion detection, PIR, and CO_2 sensors are commonly used in dynamic HVAC operation to maintain thermal comfort satisfactory level and avoid excessive energy consumption. However, none of these devices provides additional information about occupants detected(human or non-human occupants, how many occupants are there in the room?), which result in poor thermal comfort level, high or prone to a false alarm in the presence of pets or stationary object in the area of interest.

Camera-based image and video processing techniques provide additional information about occupants, which makes it more reliable to human occupants using a computer vision library template and machine learning. At the same time, the Camera-based approach does not provide an accurate number of occupants in a large capacity building during occupant movement and multiple occupancies overlapping during the period of entrance or exit from an indoor environment, which affects ventilation required to serve occupants in the indoor environments. Another challenge with the Camera-based approach is a limited range of sighting and concern of privacy.

Furthermore, other pieces of literature proposed speech processing using an Acoustic sensor to differentiate between human occupants and non-human occupants. The experimental results analysis in both works of the literature shows background sound produces poor results due to the sensitivity of the sensor to pick up the external sound even with applied background noise cancellation algorithm.

3. Future research direction

-Reduce excess energy consumption through building design: improving the current building envelop through the accounting of energy efficiency in the building design to eliminate the issue of heat loss resulting in excess heating or heat gain, which results in extra cooling. Efficient quality insulation on the attic, walls, doors, and windows should be accounted for, which prevents heat gain or loss.

-Standard pricing scheme: Another research opportunity in the smart home energy management system is the introduction of an energy pricing scheme based on energy demand from customers to reduce irregularity in energy consumption bills. The current smart home energy management system cannot provide actual pricing for energy for a day or a specific time interval, even with communication with energy providers ahead of energy usage. For example, historical energy consumption data and other dependent variables can be used to generate a model that predicts energy cost in Watt per minute to pump water on storage tankers at a particular period and provide rescheduling options to

shift the energy demand when the price is at low cost.

-Energy-saving through occupant's real-time data: One of the efficient approaches to reduce unnecessary energy consumption is to stop energy usage on unoccupied space. This can be resolve with human occupants detection procedures to differentiate between humans and animals in the space to reduce false results caused by an animal that might increase energy consumption.

-Energy-saving through low temperature turning mode: With today's reliance on technology, there is a need for self leaning and an adaptive smart home energy control system with more reliable and accurate predictions through residents'whereabouts tracking and knowledge of residents' energy usage patterns. This will help start HVAC equipment at low temperature for some time ahead of resident arrival at home to stabilize the room temperature to a desired one, which in turn saves more energy than stating HVAC equipment at an aggressive temperature setpoint upon the resident arrival.

-Thermal comfort satisfaction: The thermal comfort satisfaction of residents is one of the primary goals of the smart home energy management system. However, recent research studies are extensively focusing on eliminating unnecessary energy usage without given enough attention to thermal satisfaction of residents.

There are ways to efficiently ensure the thermal comfort of the residents and provide a healthy living environment with higher productivity.

-Improve occupants' thermal comfort through smart sensing and intelligent control: The thermal comfort satisfaction of occupants is one of the primary goals of the SHEMS. However, recent researches are constantly and extensively focusing on energy-saving without much-given emphasis on occupant's thermal comfort. Current Smart home sensing technologies are seeking new intelligent learning algorithms to promote HVAC occupants responsive research under various active states on occupant's skin temperature utilizing the wearable device and contactless sensing to infer personalized and adaptive thermal comfort.

References

[1] Simsek, Y., et al., *An analysis of additional energy requirement to meet the sustainable development goals.* Journal of Cleaner Production, 2020. 272.

[2] Hassan Qudrat-Ullah , M.A., *Dynamics of Energy, the Environment and Economy: A Sustainability Perspective.* 2020.

[3] Alrashed, F. and M. Asif, *Analysis of critical climate-related factors for the application of zero-energy homes in Saudi Arabia.* Renewable and Sustainable Energy Reviews, 2015. 41: p. 1395-1403.

[4] Asif, M., *Growth and sustainability trends in the buildings sector in the GCC region with particular reference to the KSA and UAE.* Renewable and Sustainable Energy Reviews, 2016. 55: p. 1267-1273.

[5] Iqbal, A., et al., *Interoperable Internet-of-Things platform for smart home system using Web-of-Objects and cloud.* Sustainable Cities and Society, 2018. 38: p. 636-646.

[6] Akkaya, K., et al., *IoT-based occupancy monitoring techniques for energy-efficient smart buildings*, in *2015 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. 2015. p. 58-63.

[7] Sharifi, A.H. and P. Maghouli, *Energy management of smart homes equipped with energy storage systems considering the PAR index based on real-time pricing.* Sustainable Cities and Society, 2019. 45: p. 579-587.

[8] Asif, S., et al., *Energy Management in Residential Area using Genetic and Strawberry Algorithm*, in *Advances in Network-Based Information Systems*. 2018. p. 165-176.

[9] Jung, W. and F. Jazizadeh, *Human-in-the-loop HVAC operations: A quantitative review on occupancy, comfort, and energy-efficiency*

dimensions. Applied Energy, 2019. 239: p. 1471-1508.

[10] Yang, B., et al., *Non-invasive (non-contact) measurements of human thermal physiology signals and thermal comfort/discomfort poses -A review.* Energy and Buildings, 2020. 224.

[11] Ciabattoni, L., et al. *IoT based indoor personal comfort levels monitoring*. in *Consumer Electronics (ICCE), 2016 IEEE International Conference on*. 2016. IEEE.

[12] Luo, M., et al., *The dynamics of thermal comfort expectations: The problem, challenge and implication.* Building and Environment, 2016. 95: p. 322-329.

[13] Akadiri, P.O., E.A. Chinyio, and P.O. Olomolaiye, *Design of A Sustainable Building: A Conceptual Framework for Implementing Sustainability in the Building Sector.* Buildings, 2012. 2(2): p. 126-152.

[14] Qurat ul, A., et al., *IoT Operating System Based Fuzzy Inference System for Home Energy Management System in Smart Buildings.* Sensors (Basel), 2018. 18(9).

[15] Maschi, L.F.C., et al., *Data Summarization in the Node by Parameters (DSNP): Local Data Fusion in an IoT Environment.* Sensors (Basel), 2018. 18(3).

[16] Zhang, Z., et al., *Optimal scheduling model for smart home energy management system based on the fusion algorithm of harmony search algorithm and particle swarm optimization algorithm.* Science and Technology for the Built Environment, 2019. 26(1): p. 42-51.

[17] Rodriguez, C.M., J.M. Medina, and A. Pinzón, *Thermal Comfort and Satisfaction in the Context of Social Housing: Case Study in Bogotá, Colombia.* Journal of Construction in Developing Countries, 2019. 24(1): p. 101-124.

[18] Ferrero, R., et al., *Smart Management Energy Systems in Industry 4.0.* Energies, 2020. 13(2).

[19] Hassan, N., et al., *Impact of Scheduling Flexibility on Demand Profile Flatness and User Inconvenience in Residential Smart Grid System.* Energies, 2013. **6**(12): p. 6608-6635.

Smoking Detection Algorithm using Indoor Air Quality Sensor

Yoosin Kim*, Jeajin Lee, Hyun-young Kim
Alticast Inc. Seoul, South Korea
[e-mail: yoosin@alticast.com]
*Corresponding author: Yoosin Kim

Abstract

This study introduces Smoking Detection Algorithm by collecting Indoor Air Quality such as PM2.5, PM10, TVOC in real time using cost effective sensors. The proposed model analyses and classifies the non-smoking air quality and air quality at the time of smoking by utilizing the real data from service vehicles. Further, statistically significant features for smoking detection are extracted from indoor air quality at the time of smoking and non-smoking. The proposed smoke detection model is trained using Random Forest, Decision Tree, SVM, XGBoost algorithms and the model is optimized using Gridsearch and cross validation algorithms. We also investigate performance of non-smoking/smoking classification model, cigarette/e-cigarette classification model in the context of accuracy, precision, recall, and F1-Score. All four algorithms show effective performance with the average accuracy of more then 95% for the non-smoking/smoking classification model.

***Keywords*:** Big Data, Indoor Air Quality, IOT, Machine Learning, Smoking Detection

1. Introduction

Vehicle sharing services represented by car sharing are rapidly increasing worldwide. Since the car-sharing service was first started in Switzerland in 1987, the market size has been steadily increasing, and in 2015, the market size is about 1.1 billion dollars. In 2024, it is expected to increase rapidly to 6.5 billion dollars, and the domestic market is expected to reach 500 billion won in 2020 from 100 billion won in 2016.[1]

This study collected smoking and non-smoking air quality data through an air quality measurement sensor installed in a vehicle in order to develop a model that effectively detects smoking in a vehicle. In addition, a model for classifying smoking and non-smoking air quality was developed and verified using a machine learning algorithm. In addition, we developed a model to classify the smoking air quality of general cigarettes and e-cigarettes using smoking air quality data to know what kind of cigarettes were smoked at the time of smoking. In addition, in-vehicle smoking experiments were additionally conducted to demonstrate the previously developed models.

2. Related Work

Recently, due to environmental problems such as fine dust, the need for improvement of outdoor air quality and indoor air quality is increasing, and related studies are being actively conducted. Panpaeng et al [2] proposed a cigarette smoke detector design to prevent smoking in non-smoking areas in buildings. Zaiedi et al[3] developed an air quality reading device that reads the air quality based on the derived rule by extracting the ratio value from the measured values of CO and H2 sensors based on fuzzy logic. A low-cost air quality detection system that can monitor pollutants and read cigarette smoke has been developed. In the study of Kodali et al [4], a system was proposed in which a sensor that measures CO and CO2, temperature and humidity was installed in a vehicle, and the

This research is supported by Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korean government (MOTIE) (Project No. 20192020101170)

measured value of the sensor measured in real time was transmitted to AWS in real time to inform the user. Sanger et al [5] designed a wireless cigarette smoke detection system using four gas sensors and raspberry pi, and conducted a study to implement a part that transmits and stores the measured values of the sensors to the server

3. Model

3.1 Model development procedure

A procedure is needed to develop a smoking detection model using air quality in a vehicle. The selection of models and algorithms used for air quality analysis, such as various sensors used for measuring existing air quality and smoking detection, is in progress. By applying the pre-processed data to the supervised learning algorithm, we develop a smoking/non-smoking air quality classification model and a cigarette/e-cigarette air quality classification model. During the model development process, each model is verified through measurement indices such as accuracy and precision, and the developed models are verified by conducting in-vehicle smoking experiments.

3.2 Data collection & preprocessing

Among the air quality collected in the experiment, a total of 1,180 data were generated through data preprocessing. As a result of labeling the collected data, smoking air quality was 162 and non-smoking air quality was 1,018. Of the non-smoking air quality, 88 cases are included in the normalization section of smoking air quality.

3.3 Basic Statistics & EDA

As a result of conducting the paired t-test analysis, normality test, and equal variance test, the p-values of each sensor of PM10, PM2.5, and TVOC were all less than 0.05. Therefore, it was found that there was a significant difference between the air quality of smoking regular cigarettes and the air quality of smoking e-cigarettes.

3.4 Detection Model Train & Test

To develop the smoking/non-smoking data classification model, the smoking/non-smoking air quality classification model was developed using the Decision Tree, Random forest, SVM, and XGBoosts models. For each model development, the pre-processed data set was trained and verified using a 10-fold cross validation method, and training data and test data were composed at a ratio of 9:1. As a result of developing a model using four algorithms for smoking/non-smoking air quality classification, all four models showed an accuracy of 95% or more. The model using the SVM algorithm showed the best accuracy of 95.7%, and the difference in accuracy from the rest of the model is a slight difference of about 0.5%. Models with good performance appeared differently according to model evaluation criteria.

The development of the cigarette / e-cigarette data classification model was conducted using smoking air quality data. The algorithm used previously used Decision Tree, Random forest, SVM, and XGBoosts, and the learning and verification proceeded in the same manner as the learning and verification process of the smoking/non-smoking air quality classification model. Among cigarette / e-cigarette air quality classification models, the SVM algorithm model showed the highest accuracy of 96.29%. When verifying the performance of the cigarette / e-cigarette air quality classification model based on precision, recall, and F1-score, the model using the SVM algorithm showed the best performance compared to other models.

4. Experiments

The performance of the classification model was measured, and the demonstration was conducted by observing the classified results by actually smoking cigarettes and e-cigarettes in the vehicle.

4.1 Empirical Data

In order to collect data, smoking was conducted in the vehicle, and in the experiment, smoking was conducted for a total of 4 times, with 2 times for cigarettes smoking and 2 times for electronic cigarettes. In addition, an experiment was conducted in which an alcoholic fragrance was sprayed twice assuming a non-smoking situation. Through the experiment, a total of 65 air quality data were collected, of which 11 were for smoking, 54 for non-smoking, and 6 for air quality when alcoholic fragrance was sprayed.

4.2 Smoking/non-smoking air quality classification model verification results

The performance of the smoking/non-smoking classification model shows that the model using SVM has the highest accuracy at 93.85%.

	DT	XGB	SVM	RF
True	59	58	61	58
False	6	7	4	7
Accuracy	0.9077	0.8923	0.9385	0.8923
			Total	65

Above result is a result including the detection result when the air quality is being restored to its original state.

	DT	XGB	SVM	RF
True	63	62	65	61
Retoration Interval	4	4	4	3
Actual False	2	3	0	4
Accuracy(adjusted)	0.9692	0.9538	1.0000	0.9385
			Total	65

4.3 Cigarette/e-cigarette air quality classification model verification results

The model using SVM, which had the best performance of accuracy, precision, and recall at the time of model development, had the highest accuracy in the demonstration, but showed an accuracy of about 96% in model verification at the time of development, but about 72% in the demonstration..

	DT	SVM	XGB	RF
True	6	8	7	6
False	5	3	4	5
Accuracy	0.5455	0.7273	0.6364	0.5455
			Total	11

5. Conclusions

The purpose of this study is to develop a model that effectively detects the smoking status and types of smoking(cigarette or e-cigarette) in a vehicle using inexpensive PM2.5, PM10, and TVOC sensors that are widely used in general. The air quality at the time of smoking was collected by smoking in the vehicle. After that, four machine learning algorithms (SVM, Decision Tree, Random Forest, XGBoost) were used to develop and verify smoking/non-smoking air quality classification models and cigarette/e-cigarette air quality classification models for each algorithm. Additionally, in-vehicle smoking was conducted to demonstrate the developed model.

As a result, the smoking/non-smoking air quality classification model showed the highest accuracy of 95.7% in the model using SVM, but the other models also showed high accuracy of over 95%. As for the cigarette/e-cigarette classification model, the model using SVM showed the highest performance with 96.2% accuracy, and also showed higher accuracy than other models in the verification process.

However, the model developed in this study has some limitations and problems such as liquid electronic cigarettes. In future research, the smoking behavior itself should be determined by reflecting the time-series characteristics of smoking, rather than judging by the air quality measured at any one time.

References

[1] 손종구 “카셰어링 공유경제 관심 속 시장 규모 지속적 증가” KISTI MARKET REPORT, 2017-01.

[2] Panpaeng Sancha,Phattharamon Phanpeang, and Ekkasit Metharak, " Cigarette smoke detectors for non-smoking areas in the building" *International Computer Science and Engineering Conference(ICSEC) 2018*

[3] M. K. Zaman M. Zaiedi, " Air Contaminants Monitoring of Carbon Monoxide and Hydrogen using Standalone Microcontroller Based System for Passive Smoker," *International Conference on Engineering Technology and Technopreneurship (ICE2T)*, 2014

[4] Ravi Kishore Kodali, Sasweth C. Rajanarayanan and Lakshmi Boppana, " IoT Based Vehicular Air Quality Monitoring System," *IEEE Region 10 Humanitarian Technology Conference*, 2019

[5] Junaidy B. Sanger, Lanny Sitanayah, Vivie D. Kumenap, " Detection System for Cigarette Smoke," *International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 2019.

Analysis of Research Trends for Anti-GPS Spoofing

Jiwon Lee, Yebin Park, Jong-ho Bang
Department of Software Convergence Seoul Women's University
Seoul, South Korea
[e-mail: gonea012@naver.com, talk2769@naver.com, bang6467@swu.ac.kr]
*Corresponding author: Jong-ho Bang

Abstract

Autonomous drone flies using GNSS-based navigation system and IMU sensor. GPS spoofing can disrupt GPS receiver of autonomous drone and kidnap it. Therefore, research on Anti-GPS spoofing is needed to counter the threat of GPS spoofing attack. In this paper, we classified GPS spoofing techniques and Anti-GPS spoofing techniques into three categories according to views of attackers and defenders, then proposed research trends and implications. We expect to contribute to development of drone anti-GPS spoofing technology.

***Keywords*:** Drone, GPS Spoofing, Anti-Spoofing, Soft Spoofing, Hard Spoofing

1. Introduction

The drone industry, which has been used for military purposes since the 1910s, is growing greatly as drones have been commercialized. As a result, the frequency of terrorist attacks and crimes using drones is increasing, and the damage is increasing.In September 2019, Saudi Arabia's crude oil production base was destroyed by drones, and there have been previous incidents in which terrorists used drones to threaten major organizations such as the White House and the Japanese prime minister's residence. Korea is also not a safe zone from drone terrorism.From 2016 to 2019, illegal drone flights caught near South Korea's nuclear power plants, and North Korean drones were found crashed in Paju, Gyeonggi Province, Baengnyeong Island in Incheon and Samcheok in Gangwon Province in 2014.The reason why drones can fly this long distance and attack targets is because they fly autonomously using GPS.

The drone controller, which is directly controlled by humans, uses a 2.4 GHz frequency, so the maximum flight distance is only 7 kilometers. GPS, on the other hand, eliminates the need to rely on the distance from the controller, allowing longer flights.Therefore, in this paper, we would like to pay attention to the drone's GPS Spoofing technology to defend against the attack of drones flying autonomously using GPS. The drone's GPS Spoofing is a technology that captures drone control by disrupting or simulating satellite signals, and is mainly used as a means of attack on drones.Unlike other drone incapacitation technologies, it is non-destructive, so it can minimize ground damage caused by drone crashes and can safely take over drones for investigation. Anti-Spoofing technology, which detects spoofing signals and protects drones from

"This research was supported by the MIST(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute for Information & communications Technology Promotion)"(2016-0-00022)

detected spoofing signals to counter damage caused by GPS spoofing attacks, is actively being studied.

Therefore, in this paper, we would like to pay attention to the drone's GPS Spoofing technology to defend against the attack of drones flying autonomously using GPS. Chapter 2 introduces the concept of drone spoofing and Chapter 3 categorizes various anti-spoofing technologies proposed in domestic and foreign papers into three areas, attack, detection, response, and total, according to the perspective of attackers and defenders, and then concludes with a conclusion in Chapter 4.

2. Anti-GPS Spoofing

Anti-GPS Spoofing is an anti-drone technology that responds to drone attacks. GPS Spoofing is a technique in which a spoofer intentionally manipulates GPS satellite signals and investigates them. GPS satellite signals consist of carrier L1 (central frequency 1575.42MHz), navigation data that provides coordinates, satellite time, speed, etc., and C/A codes that distinguish satellites for civilian purposes. GPS Spoofing manipulates navigation data. Jamming is a technique similar to GPS Spoofing. Jamming is a radio jamming technique that generates a signal with a stronger intensity than a GPS signal, causing the drone to confuse its position.

Unlike Jamming, an attack method using the characteristics of a GPS receiver that tracks strong signal strength, GPS Spoofing basically has the same principle as Jamming, but it is a technique that manipulates navigation data by synchronizing the spoofing signal to the satellite signal so that the drone moves to the position intended by Spoofer.

GPS Spoofing may be used by attackers as a tool for drone attacks, but there are aspects that can be used to safely kidnap intruders from the standpoint of defenders. **Fig. 1** is a picture in which a drone is examined by a spoofing transmitter for GPS-Spoofing signals, loses its lock on the Original Signal by GPS Satellites, and flies towards the Fake position intended by Spoofer, not its original destination real position.

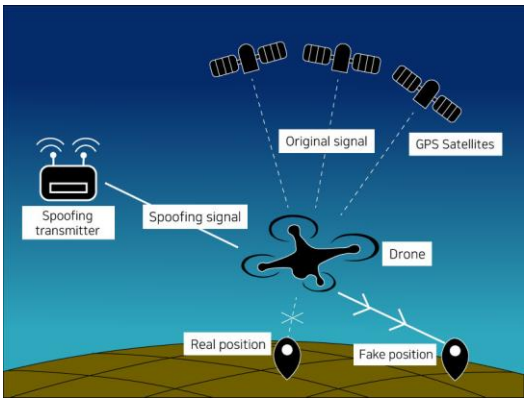


Fig. 1. Drone's GPS Spoofing

3. Classification based on the Views of Attackers and Defenders

3.1 Attack

3.1.1 Developing a new attack algorithm

The papers that studied how to kidnap drones using GPS Spoofing were classified.

. Classification of GPS Spoofing Methods [1]

- Soft GPS Spoofing: Produces spoofing signals with data similar to the navigated data of the original signal but satisfying the three synchronization conditions of the spoofing signal (strength, position offset, time offset of the spoofing signal).
- Hard GPS Spoofing: A signal that does not satisfy one of the three synchronization conditions, which can work like Jamming. Hard GPS Spoofing can be a useful striker because actual drones often don't consider loss of locks.

3.1.2 Exploring the conditions for an efficient attack

GPS Spoofing for civilian signals can be attempted with relatively low cost and effort, but the difference between satellite signals and spoofing signals makes it difficult to achieve the perfect spoofing effect. Therefore, the papers studied to increase the success rate of the attack were analyzed.

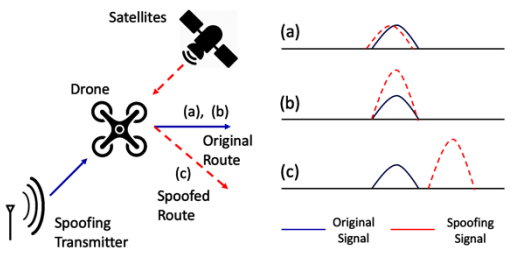


Fig. 2. Soft GPS Spoofing

. Distance from target: i.e. distance from target should be known as 22.5 m signal accuracy, i.e. the GPS receiver loses lock to the spoof signal if the spoof signal and the satellite signal have a time offset of –75 ns or more [2].

3.1.3 Presenting the weaknesses of a drone

. Drones' vulnerability classification criteria for GPS Spoofing attacks 38]

- Signal Processing Level : Vulnerable to spoof attempts through difference in signal power intensity between spoofing signal and satellite signal
- Data bit level : GPS signal frame structure is open and navigation frame information is relatively slow to change, thus vulnerable to spoofing.

. Spoofing is possible until commercial 3DR Solo drones find nine or more satellites [4].

3.2 Detection

3.2.1 Using IMU sensor value of drones

Since the IMU is an inertial navigation system that does not rely on external signals, spoofing can be detected by comparing the IMU sensor value of the drone with the acceleration value of the GPS signal [5]. If the moving speed of the drone is not constant, deceptive can be determined by comparing the acceleration measurements converted to the navigation coordinate system with the GPS estimated navigation coordinate system measurements [6].

3.2.2 Using GPS Receiver

Using Kalman filter: Adjusting the Kalman gain using adaptive fading Kalman filter minimizes the effect of the deceptive signal in a fluid manner, enabling optimal navigation to be estimated [7].

3.3 Response

3.3.1 Signal Manipulation

Spoofing signal offset through reverse phase code and carrier generation: The spoofing signal received from the drone's GPS receiver and the code and carrier waves with exact reverse phase can be generated to offset spoofing signals. However, there is a threshold that the removal of the spoofing signal is difficult if there is a synchronous error between the spoofing signal and the reverse phase code [8].

3.3.2 Developing Spoofing Response System

Safe-hijacking: Proposing a safe method of hijacking from the damage that falls while authorizing an intruding drone. "Drones in the DJI series", examining signals in the opposite direction of their actual location; "Parrot Bebop," manipulation of spoofing signals through the drone's path-finding algorithm; "3DR Solo," drones use Soft GPS Spoofing to keep them locked; [1]

3.3.3 Developing Anti-Drone System

Seven circular array antenna structures: The existing array antenna structure and other structures remove spoofing signals even in environments where location information is unclear, and normal GPS signals can be received without loss. Propose an array antenna operating scenario that efficiently beamforming and knurling through multistage mode [9].

4. Conclusions

In this paper, the papers that studied GPS Spoofing and corresponding techniques made so far were divided into three categories according to the views of attackers and defenders. In the field of attack, studies on environmental conditions for effective attacks were mainly made, indicating optimal signal power and attack distance for the target. When trying to detect GPS Spoofing from the perspective of a defender, it was possible to detect spoofing signals through IMU sensors mounted in drones, or through single receiver, multiple receiver, and artificial intelligence learning. In response techniques after spoofing signal detection, there were

techniques to offset or eliminate spoofing signals through the formation of operating signals, and techniques to suppress spoofing signals by using specific receivers or array-antenna systems. Through this paper, we were able to grasp the overall risk of drone attacks using GPS Spoofing and the level of response technology

References

[1] JUHWAN NOH, "Tractor Beam: Safe-hijacking of Consumer Drones with Adaptive GPS Spoofing", ACM, 26 pages. Article 12 April, 2019.

[2] Nils Ole Tippenhauer, "On the Requirements for Successful GPS Spoofing Attacks", ACM, 2011.

[3] Ali Jafarnia-Jahromi, "GPS Vulnerability to Spoofing Threats and a Review of Antispoofing Techniques", International Journal of Navigation and Observation, 16 pages, 2012.

[4] SANDRA PÉREZ ARTEAGA1, "Analysis of the GPS Spoofing Vulnerability in the Drone 3DR Solo", IEEE, 2019.

[5] Zhiwei Feng, "Efficient Drone Hijacking Detection using Onboard Motion Sensors", IEEE, 2017.

[6] Keum-Cheol Kwon, "A Study on GPS Spoofing Detection Method Using AHRS and Acceleration Measurements", The Graduate School Chung-Ang University, 2018.

[7] Sun Young Kim, Chang Ho Kang, Chan Gook Park, "Station Based Detection Algorithm using an Adaptive Fading Kalman Filter for Ramp Type GNSS Spoofing", Institute of Control, Robotics and Systems, 2015.

[8] Taehee Kim, Sanguk Lee, Jaehoon Kim, "The Anti-Spoofing Methods Using Code Antiphase of Spoofing Signal", The Journal of Korean Institute of Communications and Information Sciences, 2013.

[9] Jung-Su Han, Ki-Yun Kim, Seok-joong Kim, Hyun-do Kim, Hyung-Jin Choi, "Analysis of Adaptive Digital Signal Processing for Anti-Jamming GPS System", The Journal of Korean Institute of Communications and Information Sciences, 2007.

A System Structure for Detecting the Anomaly Data and Energy Consumption in the Real-time Energy Monitoring System

Yeonjee Choi*, Hyunsuk Hwang*, and Taeyeon Kim*
*Technical Research Center, New Communication Infotech Co., Ltd, Busan, South Korea
[e-mail: chbigbrother@gmail.com, hwanghs29@gmail.com, ncinf@hanmail.net]
*Corresponding author: Hyunsuk Hwang

Abstract

As the smart factorization progresses, the number of real-time monitoring systems are being developed by using IoT devices. In this paper, we build a system that provides real-time notifications to field users by determining whether real-time data collected by the IoT-based real-time energy forging monitoring system has abnormal data collection and energy consumption. This system consists of an IoT platform data collection server, a monitoring and management web server, a data analysis and prediction server, and a user service, and includes functional modules for managing anomalies. The IoT platform data collection server collects energy data from each facility in real time, transmits it to the monitoring and management web server according to the classified abnormal types of cases s, and provides a notification service to the related users. This system has an advantage on saving energy and establishing effective work plans by using deep learning technique. It provides each facility's anomaly detection events in real-time by predicting gas consumption and elapsed time and connecting with the prediction server, which detects the gas consumption anomalies.

***Keywords*:** Real-time Energy Monitoring System, Smart Factory, Forging Process System, Energy Anomaly Detection System, Deep Learning-based Predication System

1. Introduction

The anomaly detection system in smart factory is considered as an essential element for factory automation. For factory automation, a system that can receive and monitor environmental data in real time is important. Above all, an anomaly detection system is also being developed based on real-time data. Recently, the development of IoT-based services for automation system is actively progressing[1]. According to the McKinsey survey, it is predicted to generate an economic value of 11.1 trillion dollars in IoT by 2025 [2]. A system that collects and monitors real time data by connecting the facilities used in factories and IoT devices and analyze data collection of anomalies is becoming an essential function. Furthermore, a service that predicts anomalies and provides real-time notifications to users through machine learning is being developed [3]. This can play a role of minimizing the maintenance time of abnormal conditions of machines and facilities and preventing risk factors in advance through anomaly detection. Therefore, we propose a method for constructing a system to reduce excessive energy consumption by notifying users in real-time and allowing them to response quickly to the situations.

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning(KETEP) and the Ministry of Trade, Industry & Energy(MOTIE) of Republic of Korea(No. 20182020109660).

2. Related Research

2.1 Real-Time Monitoring System

Three essential elements to build a smart factory for automation are IoT-based device technology, process and analyzing the data transmitted from factories in real time, modeling platform technology, and manufacturing application technology [4]. To process such data, recently monitoring systems collect factory's temperature, humidity, pressure, electricity based on IoT devices and transmit the data in real time. Due to this smart system of the factory, various expectation effects can be obtained such as a 15% increase in productivity and quality improvement [5].

2.2 Anomaly Detection System

Research on existing smart factory anomaly detection systems are being conducted through AI machine learning methods. To analyze and detect data anomalies in time series, an LSTM model was used, which is based on an approach using big data [6]. In addition, research such as improving manufacturing quality by predicting future data and errors through learning anomaly patterns using collected data using industrial IoT devices are being conducted [3]. This is because data analysis through machine learning is becoming more common, and it is possible to supplement or replace abnormality detection by using machine learning in an environment where a large amount of data is collected. The IoT platform, which the part is applied, is useful in applying anomaly detection because it can transmit and receive real-time data in a short period. Rather than implementing a predictive model, this study focuses on the construction of real-time anomaly data collection and transmission system used in forging factory.

3. Structure of Data Anomaly Detection System

3.1 IoT-based System Structure

The important modules of the structure of the anomaly detection system using IoT platform presented in this paper are composed of three modules. Data transmission interface module for IoT server, anomaly collection management and notification module in real-time monitoring server, and the prediction model of the prediction server using machine learning and anomaly detection module. **Fig. 1** shows the structure of data anomaly detection system. The IoT platform was used to collect energy consumption data per minute from the factory in real time. In order to collect abnormal data by error type, it is mainly classified into hardware, software, and physical power errors. For example, the physical (hardware) symptom is HW, the software symptom is SW, and the power-related symptom is PW. In addition, for the detailed errors related to these main categories can be classified by serial number. Based on these error codes, they are transmitted through REST API from the IoT platform. Regarding transmitted errors, each error code is stored in a database with its own details and provides notification to users and history display function through error collection management module. In the prediction server, it displays actual predicted process time and usage data using the constructed LSTM learning model by receiving the time-series data. The prediction model determines anomalies depending on whether cumulative elapsed time and the cumulative usage decreases or the data is not collected over time then transmits them to the error collection management module. Finally, users can receive notifications of energy accumulations, temperature anomaly collection, process time and usage anomaly from the system.

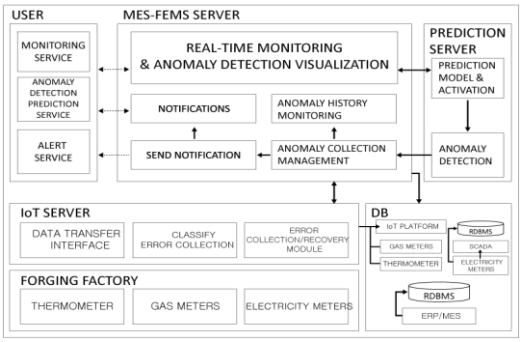


Fig. 1. System Structure and Function Modules

3.2 Anomaly Notification Structure

Fig. 2 shows error code structure for sending anomaly notification based on the IoT server.

ERROR CODES

PW - POWER	001 IoT Server Power Disconnection
	002 Furnance Gateway Power Off
	003 Heat Treatment Gateway Power Off
	004 Data Collecting Housing #1 Power Off
	005 Data Collecting Housing #2 Power Off
	006 Data Collecting Housing #3 Power Off
	007 Data Collecting Housing #4 Power Off
	008 Data Collecting Housing #5 Power Off
SW - SOFTWARE	001 Reboot after Windows Update
	002 SSH Remote Access Disconnection
	003 Furnance Gateway Wireless Network SW Error
	004 Heat Treatment Gateway Wireless Network SW Error
	005 Furnance Gateway Data Transmission SW Error
	006 Heat Treatment Gateway Data Transmission SW Error
	007 RaspberryPie Wireless Network SW Error
	008 Gas Data Collection SW Error
HW - HARDWARE	001 Furnance Gateway Anomaly
	002 Heat Treatment Gateway Anomaly
	003 RaspberryPie Anomaly
	004 Gas Compensator Anomaly
	005 Low Gas Compensator Battery
	006 Thermostat Anomaly
	007 Power Meter Anomaly
	008 Data Network Line & Connector Anomaly
	009 Internet Disconnection
ETC	Other Errors

Fig. 2. Classified Error Codes for Anomaly Data Collection

The IoT server identifies anomalies based on real-time data transmitted from the forging factory. Anomalies are delivered to the monitoring server along with the detailed explanations. At this time, not only the anomaly data is transmitted but also the recovery data is delivered to the system when the automatic recovery is happened. Monitoring server has error codes that are classified by different cases. When anomaly data is transmitted from IoT device, it matches with the existing error codes and stores the data in the database, at the same time and sends an e-mail to the error-related workers by e-mail. The prediction server transmits a predicted anomaly values to the users, which is predicted by the prediction model using real-time data delivered every 2 minutes in 10-minute increments. In prediction server, apart from the IoT server's anomaly notification, can give two different information with the anomaly detection. First, during work, the prediction model predicts the estimated working time and when the actual work doesn't ended at the time that is predicted, it gives notification to the worker. Second, the notification is related to the gas usage and furnances' temperature. It can predict the working and heating end time that it compares the real time working data and the data has a huge difference between the two, e-mails are sent so that operators or emergency contacts can be aware of the equipment collection abnormality.

4. System Implementation

4.1 Anomaly Data Collection Structure

In order to collect anomaly data from IoT platform, JSON formatted value is delivered through asynchronous method through REST API. **Fig. 3** is a database table used for facility collection and management. In the transmitted data, not only the error code but also the status name and error occurrence time, cause, solution, facility name, notification sender, and possibility of automatic recovery are included. All received data is stored in the database and used as the data for further improvement of facilities and make statistics on each error code's frequency of occurrence. Based on the received data, an alarm is sent with e-mail to multiple users using SMTP (Simple Mail Transfer Protocol). At the same time, it stores the details of error to the database so that the error history can be searched, and statistics are generated to imporve recovery time on frequent errors in the future. In the forging factory, there are the number of heating furnaces, heat treatment furnaces, and presses, which are then classified by being connected to three types of enclosures in our experiment enviroment. To dynamically manage all types of equipment and error information, a total of three tables were created in the database. Each of the three tables stores error codes, collects, and stores equipment collection data, and stores necessary information when viewing notifications. In error occurrence management, errors except for software problems are recorded directly on the error inquiry page and the recovery method and action taken.

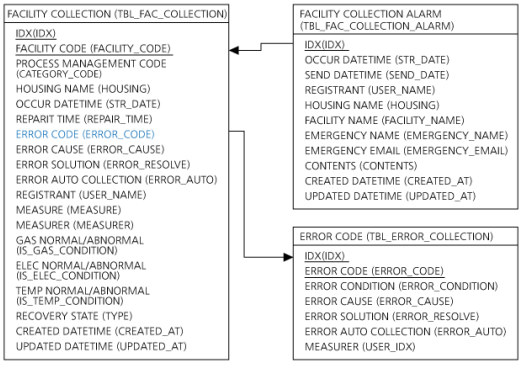


Fig. 3. Facility Collection Management Table Structure

4.2 Implementation

1) Facility Collection Status Management

Fig. 4 is a screen of facility collection status management for collecting anomaly data. On the facility collection status page, users can check the exact values such as temperature value and gas volume and the cause of each error through the error detailed information in the error list chart.

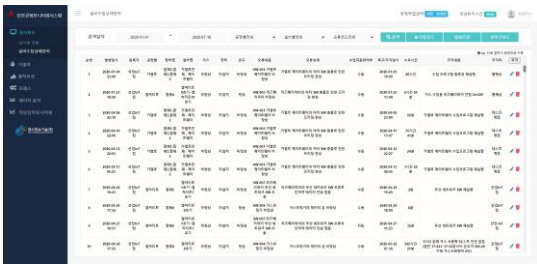


Fig. 4. Facility Collection Management Table Structure

In addition, you can check the status name for each error, the time of occurrence of the error, the cause, the solution, the name of the facility, the sender of the notification, the automatic recovery status, the recovery time, the recovery status, and the recoverer. By collecting these data, on the monitoring visualization page, it analyzes the time demands depending on each error code. This is for shortening the time and effort to recover the anomalies and improve the accuracy of anomaly detection in the future. However, due to the current automated system, problems such as physical or power off need to be manually recorded by the operator after taking measures, that history management is not well performed. Therefore, it is necessary to classify possible actions for errors each person in charge of the field site. The person in charge is classified by error code, and the alarm can be sent to workers who can recover the error. Also, it gives the information about the specific facility that it is possible to notify the registered worker about the error. Regarding this anomaly detection notification is giving as a real time data, this is an important part. Because when the notification gives every worker in the field, the importance of it can be reduced and the chance to ignore the error that they are involved is increasing.

2) Visualizing Anomaly Data History

Above all, utilizing the advantage of the IoT platform enables quick response and recovery by allowing users to visually check real-time error conditions on the history search display page that displays updated data every minute in a chart. The prediction server consists of a prediction model and an anomaly detection module. Conduct a machine learning in a virtual environment using Tensorflow installed in Docker's container. The trained model receives actual data in real time through REST API and transmits the value in JSON format. The transmitted data is stored in the database. Stored data will be displayed seperately in another menu to statisticalize the history of anomaly prediction system. Through this list the system, we are expecting that the system can be improved by classifying the various errors that can happened during the work. At this time, the anomaly detection data that the worker wants to know is an important part in the process work. Gas or temperature anomalies in a heating furnace can have a great effect on saving energy in a plant. In addition, by notifying the operator of the completion time of the process operation in advance through the predicted data, the next operation can be prepared in time, and unnecessary energy consumption can be prevented in the event of a facility's anomaly. **Fig. 5** is a screen showing temperature anomaly detection information on the history search screen. This visualized history can apply to the prediction visualization system and show users the predicting working and heating duration and end time by receiving the real time data.

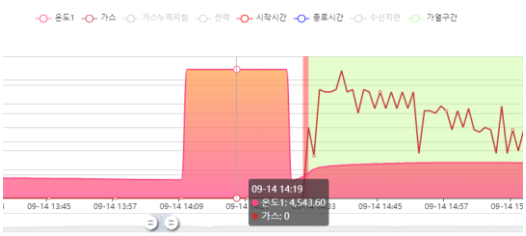


Fig. 5. History Search Page

5. Conclusions

This paper proposes a system structure of anomaly detection data collections based on real-time data and implementation of

visualization program through IoT device and machine learning for rapid reponse and facilities' anomaly configuration. When an anomaly occurs in the real-time data, the anomaly data, which is classified by the error phenomenon, is received and a notification is given to the person in charge. The advantage of this structure is the function of giving immediate notification to the person in charge when anomaly is detected, especially when it detects the anomaly in gas or temperature consumption, and estimated working elapsed times are predicted during forging process work through machine learning. This makes it possible for the workers to immediately respond to the error situation that proceeds in real time and prevents double work by sending a notification about errors, which is automatically recovered or not. In the future, it is necessary to analyze the collected error data to find out and improve the program whether there is aging equipment, which work processes need to be improved, and what are the same errors that occur redundantly. In addition, it is possible to provide a way to speed up the recovery time by analyzing how long it took to recover. However, if an anomaly notification is not efficiently delivered to worker or working progress status is not checked by the automated system, it can be overlooked even though an error has occurred due to not paying attention. The issue of designating a person in charge of a task and an unfamiliar notification delivery method can be an important issue for efficient notification delivery. Therefore, continuous updates with the working process is important and negotiating the other possible errors that can happen frequently is required for the internal staffs. In addition, in order to provide a more accurate anomaly notification function, the system will be implemented into an actual program and apply the prediction model and facility anomaly detection. With test result of the categorization of errors and anomaly detection notification system, it is expected to be adapted into the future machine learning based anomaly detection monitoring system, which will have higher accuracy in predicting anomalies.

References

[1] G.M. Bae, S.G. Jang, M.S. Kim, I.W. Joe, "Autoencoder-Based on Anomaly Detection with Intrusion Scoring for Smart Factory Environments," (eds) *Parallel and Distributed Computing, Applications and Technologies*, PDCAT 2018. Communications in Computer and Information Science, Vol 931, 2019.

[2] Ménard. "How can we recognize the real power of the Internet of Things? How Can We Recognize the Real Power of the Internet of Things? ," https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/how-can-we-recognize-the-real-power-of-the-internet-of-things#, 2017

[3] S.Y. Im, H.K. Choi, K.Y. Yi, T.H. Lee, and H.C. Yu, "Anomaly Detection System of IoT Platform using Machine Learning," *Proceedings of the Korea Information Processing Society Conference*, pp. 1001-1004, 2018

[4] H.S. Hwang, Y.W. Seo, T.Y. Kim, "The Development of a Web-based Realtime Monitoring System for Facility Energy Uses in Forging Processes". *Journal of Internet Computing and Services*, Vol 19. No.1, pp. 87-95, 2018.

[5] S.H. Choi, "Proposal for data collection and systematization for smart factory construction". *Advantech* pp.114-115, 2017.

[6] V.Q. Nguyen, L.V. Ma, J.S. Kim, "LSTM-based Anomaly Detection on Big Data for Smart Factory Monitoring". *Journal of Digital Contents Society*, Vol 19. No.4, pp.789-799, 2018.

Consumer Preference for Electricity Pricing's Choice Rate System based on IoT : The case of South Korea

Solji Nam[1], Changeun Park[2] and Jungwoo Shin[1*]
[1]Department of Industrial and Management Systems Engineering, Kyung Hee University, South Korea
[2]Department of Applied Bigdata systems Engineering, Kyung Hee University, South Korea
[e-mail: sjnam@khu.ac.kr, anyplepark@khu.ac.kr, shinjung11@khu.ac.kr]
*Corresponding author: Jungwoo Shin

Abstract

This paper investigates consumer preferences on the introduction of choice rate system for electricity pricing. Due to the unique power structure in Korea, the variety of rate plans is relatively low. However, now the power generation sources are diversified and run on smart grids. Therefore, a new rate system that can manage power demand using IoT technology in real time is needed. The proposed optional rate system allows you to choose the time to contract the power plan and the amount of power used in that time zone, and to choose the percentage of energy sources to be supplied. In addition, social costs is also considered. Five main attributes are used to analyze consumer preference and the survey was conducted at each level. According to the analysis, the smaller the contract power usage and the lower the social cost, the more preferred they are. In addition, the higher the supply ratio of renewable energy, the more preferred it is. In Korea, renewable energy-related factors need to be included in the electricity bill scheme.

***Keywords*:** Electricity Pricing, Choice Rate System, Choice Experiment, Internet of Things (IoT), Social Cost, Smart Grid

1. Introduction

In 2019, Korea Electric Power Corporation (KEPCO) recorded the second-largest deficit ever. This is because profits from electricity sales have decreased and operating costs due to greenhouse gas regulations have increased significantly.[1,2] Furthermore, electricity rates should be of interest as well as profitability, so it is important to establish a reasonable rate system.

For example, the government is considering a new 'Green Pricing' that will voluntarily pay an additional amount to existing electricity bills to increase demand for renewable energy. This is a new perspective, and while the policy has been targeted at power generation companies, Green Pricing is targeted at power consumers.

In addition, the Ministry of Trade, Industry and Energy is conducting a future smart grid demonstration research project in July 2019. This is currently being discussed by SK Telecom Consortium and Gwangju to plan the optimal demonstration complex. In the end, the reason why smart grids can be applied is that due to the development of IoT technology, real-time

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2017H1D8A2031138).

power demand can be managed. Furthermore, each consumer's big data management will allow them to provide optimal usage services.

Currently, Korea has a special power structure, and the electricity bill itself is low, so there is a limit to the fact that there is not much incentive to choose various plans. In particular, although the time-based rate system is being applied to consumers for industrial and general use, the government is still preparing to introduce a time-based rate system as an alternative to the progressive system. Therefore, a study of consumer preferences for the diversity of power charges is needed.

2. Method

2.1 Choice Experiment

This study uses a discrete selection model based on random utility theory that analyzes consumer utility for multiple properties.
The most widely used discrete selection model i s a logit model, expressed in closed form to facil itate calculation. The discrete probability equatio n is as follows.

$$P(Y_n = 1) = \frac{\exp(\beta' x_n)}{1+\exp(\beta' x_n)}$$

The probability of selection is between 0 and 1, a nd the sum of the probability of selection for all a lternatives is 1.

2.2 Data and Survey Design

The study used the Stated Preference (SP) data to analyze consumer preferences for virtual market conditions called optional plans. The data was collected from various parts of Korea to reflect the population distribution of all people. The survey is currently conducted until a pilot survey, with respondents aged between 20 and 60 years old. The pilot survey was conducted online from September 5 to September 9, 2020. It was done by research firm EMBRAIN PUBLIC. The study distribution among respondents surveyed is as follows:

Table 1. Demographic properties of sample.

Category	Characteristic	Respondents(n)	Percentage(%)
Total		100	100.0
Gender	Male	51	51.0%
	Female	49	49.0%
Age	19-28	18	18.0%
	29-38	18	18.0%
	39-48	23	23.0%
	49-58	23	23.0%
	59-68	18	23.0%

The pilot survey conducted a concoction method to identify the consumer preference structure of the optional power rate system. The concoction approach is a way to present a combination of elements with virtual characteristics to consumers, and to order preferences to measure consumer preferences.[3] The research suggests a virtual selective rate system and performs analysis using the results of selecting respondents. The following five attributes were discussed to analyze consumer preferences of the optional rate system: (1) Power usage fixed time zone, (2) Contract power usage, (3) Energy source selection status, (4) Energy source supply ratio, (5) Social cost [4-6]

Table 2. Attribute and attribute levels

Attribute	Level
Power usage fixed time zone	Current During maximum peak exclusion time During maximum peak load time For a day
Contract power usage	200kWh / 300kWh

	100kWh / 200 kWh
	100kWh/ 300kWh / 500kWh
Energy source selection status	Current
	Selectable energy source
Energy source supply ratio	30% : 40% : 20% : 10% 25% : 30% : 25% : 20% 20% : 20% : 30% : 30% 15% : 10% : 35% : 40%
Social cost	Not reflected 10% of Electricity pricing 20% of Electricity pricing

The case number of 8x5x3=120 is necessary to represent all the characteristics and levels set up in this study. Since the number of all cases cannot be presented to respondents in reality, this study reduced the total number of alternatives to 52 by forming an orthogonal matrix. Five alternatives (including nochoice) were then grouped into one set to respond to respondents' preferences. Five of the 13 choice sets were presented to the respondents at random.

3. Analysis

Table 3. Estimation result of logit model for Choice rate system (Base model)

Attribute	Mean of β	Variance of β
Power usage fixed time zone	0.0686	0.0548
Contract power usage	-0.0008***	0.0004***
Energy source selection status	-0.0405	0.2694
Energy source supply ratio	0.0039*	0.0102*
Social cost	-0.1671***	0.0008***

Note: *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

The analysis results are shown in **Table 3**. It was not noted for power usage fixed time zones. The smaller the contract power usage, the lower the social cost, the more preferred it is. In addition, the higher the proportion of renewable energy, the more preferred the energy supply.

4. Conclusions

This study is an analysis of consumer preferences for the optional rate system of power charges. Using IoT technology, real-time demand is managed, enabling strategic electricity use. This means that it could be a catalyst for consumers' preference for choice rate system. This study showed that there was a high interest in the optional rate system with a high supply ratio of renewable energy. Through this, it is necessary for Korea to include renewable energy-related sectors in its power rate components.

References

[1] Y. S. Lee, and H. J. Park, “An Analysis of the Consumer Welfare Changes in the Selective Power Rate System,”, *Korea energy economics institute*, 2013.

[2] B. S. Choi, and S. H. Yang, “A Study on Efficient Power Management of Home Energy Storage System Under Time of Use Electricity Pricing,” *The Korean Institute of Power Electronics*, pp.228-231, 2019

[3] P. E. Green, and V. Srinivasan, “Conjoint analysis in consumer research: issues and outlook,” *Journal of consumer research.* Vol 5. No. 2, pp. 103-123, 1978.

[4] A. Jongejan, B. Katzman, L. Thomas, and M. Michelin, “Dynamic Pricing Tariffs for DTE’s Residential Electricity Customers,” *Center for Sustainable Systems*, Vol 00, No. 0, pp. 0-00, 2010.

[5] A. Faruqui, R. Hledik, and J. Palmer, “Time-Varying and Dynamic Rate Design,” *The Regulatory Assistance Project*, Vol 00. No.0, pp. 00-00, 2012.

[6] H. J. Welinghoff, and D. L. Morenoff, “Recognizing the importance of demand response : The second half of the wholesale electric market equation,” *Energy Law Journal*, 2007.

Analysis of Compressed Sensing Method and Performance of Large IoT Sensor Data

Hye-min Kwon, Young-woo Lee, Min-jun Kim, Chang-kyo Lee, Jong-han Choi, Min-goo Kang, Jeongwook Seo
Namseoul University, Gumi Electronics & Information Technology Research Institute,
Hanshin University
Seoul, South Korea
[e-mail: jwseo@hs.ac.kr]
*Corresponding author: Jeongwook Seo

Abstract

As IoT(Internet of Things) sensors and collection technology advance, research is being actively conducted to derive meaningful results of large data. However, because the shape and quality of data are different depending on the characteristics of the IoT sensor and the collection environment, a preprocessing process is required to obtain refined data. In addition, in order to collect large data from IoT sensors in real time, efficient feature extraction technology that can reduce the size of data is required. Therefore in this paper, we analyze the compressed sensing method and reconstruction performance that extract features from image data collected from camera sensors installed in IoT devices. It was confirmed that the size of data decreased according to the compression rate of the image data, and the corresponding restoration performance was compared. Also, the AP(Average Precision) of the YOLO model was confirmed.

***Keywords*:** Compressed Sensing, Large Sensor Data, IoT(Internet of Things), Average Precision

1. Introduction

With the development of IoT Sensing Technology, large data is generated and with the development of dig data processing systems, large sensor data analysis and efficient use are being conducted [1]. However, for large data, the configuration and size of the input data are variable, and the shape and quality of the data are different depending on the IoT sensor used, so a preprocessing process to obtain refined data is required [2].

Preprocessing refers to the process of classifiying, removing, and classifying data unnecessary for learning and data essential for learning before learning, which can improve various performance [3].

In particular, when using large data, and efficient feature extraction technology that can reduce the size of data is required. Among them, the compressed sensing used in this paper is a new signal acquistion theory that can accurately restore a signal sampled below the Nyquist rate to the original signal when the signal has a sparse characteristic and the linearly measured values are incoherent [4,5].

Therefore, in this paper, we analyze the compressed sensing method to extract features from image data, the reconstruction performance, and the performance of the object detection model YOLO.

2. Compressed Sensing Method of Large IoT Sensor Data

2.1 Compressed Sensing

The most interesting part of compressed sensing theory deals with the case where the signal can be sent completely without sampling above the Nyquist rate [6]. According to the compressed sensing theory, the original signal can be almost restored with very little linear measurements [7]. For compressed sensing, installed Ubuntu 18.04.4 and python 3.8 or higher. The compression rate was 50%, 40%, 30%, 20%, 10% and 1000 images were compressed sensing for each compression rate. The figure below shows the restored image according to the compression rate and the original image.

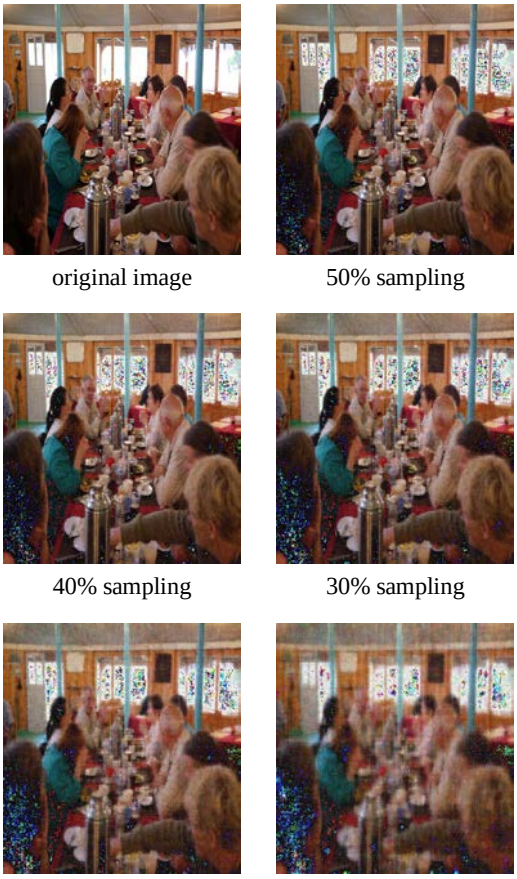

original image
50% sampling
40% sampling
30% sampling
20% sampling
10% sampling

Fig. 1. Compressed Sensing images

2.2 Object Detection Model : YOLO

YOLO stands for You Look Only Once. It is an algorithm that estimates the type and location of an object by looking at the image once by considering the bounding box and class probability of an image as a single regression problem. The figure below shows the result of object detection using the YOLOv5 model for the reconstructed image and the original image after compressed sensing at compression rates of 50%, 40%, 30%, 20% and 10%.

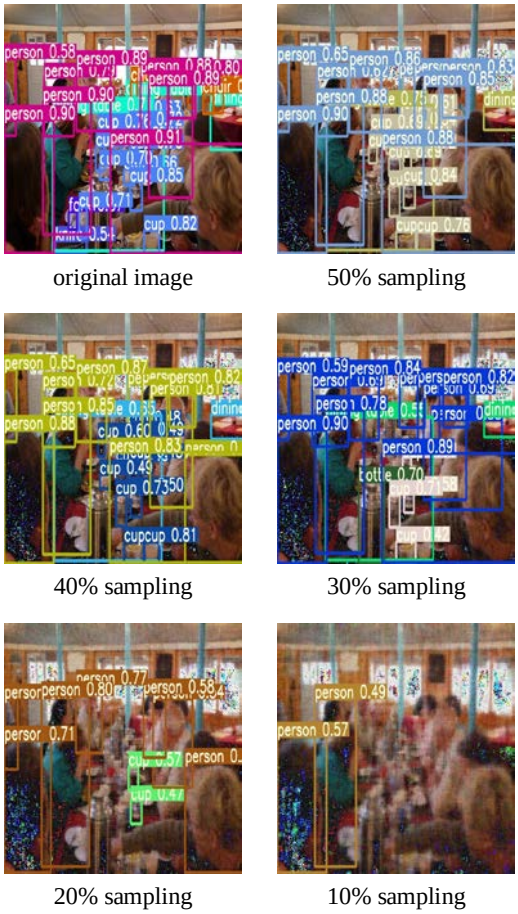

original image
50% sampling
40% sampling
30% sampling
20% sampling
10% sampling

Fig. 2. Object Detection: YOLOv5 images

3. Restoration Performance Comparison

In this paper, compressed sensing images were compared using YOLOv3 and YOLOv5 to compare the reconstruction performance according to the compression rate. The accuracy of object detection was indicated by (the number of objects detected in the compressed sensing image/ the number of objects detected in the original image). AP was compared for each version of the YOLO model. AP is an index used as an evaluation criterion for PASCAL VOC. The performance of the object detection algorithm can be expressed as one index. The

closer the AP is to 1, the higher the performance of the object detection algorithm [8]. The figure below is a graph showing the restoration performance according to the compression rate when the YOLOv5 and YOLOv3 models are used. The red graph show the restoration performance of YOLOv5 and the blue graph shows YOLOv3.

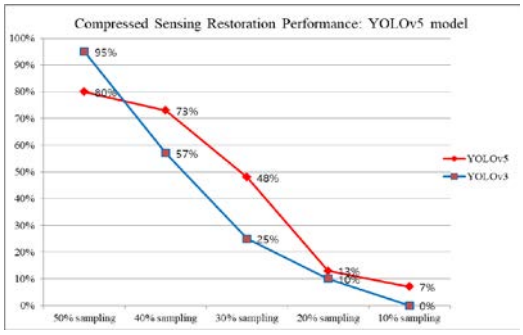


Fig. 3. Compressed Sensing Restoration Performance: YOLO model

It can be seen that the restoration performance of the YOLOv5 model is better than that of the YOLOv3 model, and the AP of the YOLOv5 model is as follows.

```
Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.388
Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.562
Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.411
Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = 0.143
Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = 0.436
Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.633
```

Fig. 4. AP of the YOLOv5

4. Conclusions

In this paper, we analyze the compressed sensing method and performance of large IoT sensor data. After compressed sensing 1000 original images at 50%, 40%, 30%, 20% and 10% compression rates, the reconstructed images were evaluated using the YOLO model, and object detection model. Depending on the compression rate, when using the YOLOv5 model, the accuracy of object detection was confirmed at about 80%, 73%, 48%, 13% and 7% and when using the YOLOv3 model, about 95%, 57%, 25%, 10% and 0% accuracy was confirmed. Therefore, it was confirmed that if the image data is compressed sensing by 50% sampling, the size of data can be greatly reduced, and when object detection is performed compared to the original image, almost similar results can be obtained. In the future, we will study not only image data but also compressed sensing of various large IoT sensor data.

References

[1] Jieun Kim, Hana Lee and Kyungchang Kim, “Database Design for Efficient Processing of Large-Scale Sensor Data,” *The Korean Institute of Information Scientists and Engineers*, pp. 244-246, 2015.

[2] Yoosang Park, Jongsun Choi and Jaeyoung Choi, “A Method to Provide Context from Massive Data Processing in Context-Aware System,” *Korea Information Processing Society*, Vol 8. No.4, pp. 145-152, 2019.

[3] Donghyun Kim, Seungeon Yoo, Byungjun Lee, Kyungae Kim and Heeyoung Youn, “Data preprocessing for efficient machine learning,” in *Korean Society of Computer Information*, Vol 27. No.1, pp. 145-152, 2019.

[4] Thu L. N. Nguyen, Honggyu Jung and Yoan Shin, “A Signal Detection and Estimation Method Based on Compressive Sensing,” in *The Journal of Korean Institute of Communications and Information Sciences*, Vol 40. No.6, pp. 1024-1031, 2015.

[5] Junwon Choi, Byonghyo Shim, Yacong Ding, Bhaskar Rao and Dongin Kim, “Compressed Sensing for Wireless Communications : Useful Tips and Tricks,” in *IEEE Communications Surveys & Tutorials*, Vol 19. No.3, pp. 1527-1550, 2017.

[6] Heungno Lee, Sangjun Park and Suncheol Park, “Introduction to Compressed Sensing,” in *The Institute of Electronics and Information Engineers*, Vol 38. No.1, pp. 19-30, 2011.

[7] Samina Khalid, Tehmina Khalil and Shamila Nasreen, “A Survey of Feature Selection and Feature Extraction Techniques in Machine Learning,” in *Science and Information Conference*, 2014.

[8] Sangyoon Park, Sanghyun Yoon and Joon Heo, “Image-Based Automatic Detection of Construction Helmets Using R-FCN and Transfer Learning,” *Korean Society of Civil Engineers*, Vol 39. No.3, pp. 399-407, 2019.

Security Challenges of Blockchain in Digital Healthcare and Internet of Things Systems: A Review

Aitizaz Ali[1], Muhammad Fermi Pasha[2], Muhammad Hasnain[3], and Rafia Afzal[4]
[1,2,3] School of Information Technology, Monash University Malaysia.
[aitizaz.ali@monash.edu, muhammad.FermiPasha@monash.edu, muhammad.malik1@monash.edu]
*Corresponding author: Aitizaz Ali

Abstract

The Electronic Health Record (EHR) has a notable influence on modernizing and transforming society and industry to digitize health knowledge. Digital healthcare system is undergoing exponential development in industry and research. However, it contains some security and privacy vulnerabilities. The research community has put a significant effort into blockchain technology as a new security tool. In this research, blockchain technology's considerable role and applications have been studies to identify security issues in the digital healthcare system using blockchain. In this review paper, popular electronic databases were used to find publications on the research topic. The digital healthcare system and blockchain's security mechanisms have been ranked into three main categories, smart city, data management, and information, providing architecture and framework. The findings of this research offer distinctive technical directions of blockchain that might impressively help to find a solution for privacy and security problems encountering the digital healthcare system design and supply distributed storage, transparency, trust, and other support for patient-related information. This paper's findings also support forming a valid, impressive, and secure distributed healthcare network and guarantee the privacy and security of participants involved in the healthcare system.

Keywords: Privacy, electronic health records, blockchain, smart city, attacks.

1. Introduction

Recently, there has been notable progress in industrial designs, information and communication systems, wireless communication technology, and electromechanical systems, encouraging novel technology called the Internet of Things (IoT). IoT's essential desire is to link total or specific devices to the other connected devices or the Internet [1-2]. Therefore, IoT is the connection of computing devices, which remain embedded in everyday things to transfer information or get data via the Internet [3]. The Internet considerably alters the human's living method and the communication among them from professional life to social connections. IoT makes the Internet enable relations among smart/small devices, which bring up a combined establishment of physical plus virtual objects. The potency of various IoT services and tools to send gathered data and utilization of converted data is called interoperability [3-4]. It affects dispensable services, functionality, alteration management, and the chances of restricted resource utilization [1, 5].

The opinion about networking appliances and other things is approximately new [6]. There are some challenges to privacy and security. Incorrect signal injection, spoofing, eavesdropping, and replay attacks are the fundamental attacks on IoT systems [7-8].

Such attacks negatively influence security features, including reliability, authenticity, and privacy. So, it is necessary to consider these challenges and analyze the protection of IoT products and services. Since the hack of any device might provide access to the user data, the services and data must be secured from the cyber-attacks [9]. Albeit, the IoT industry progress causes IoT security to turn to an industrial hot topic. Also, a universal concern is if IoT may provide diverse users and domains [10]. The smart communication of people, objects, data, and processes abruptly enhances connected devices throughout the globe, expected to 50 billion by 2020 [11]. Therefore, the protection complexities of IoT applications are required to deal at diverse stages, but intricacy enhances as the great device heterogeneity lacks in performance. Access control is the main trouble for any source of data. Therefore, it is essential to have legal access to confidential data for certifying integrity, privacy, confidentiality, and authenticity [12-13].

This paper aims to answer the following research questions:

RQ1: How Blockchain technology is solving the privacy and security concerns of personal healthcare record?

RQ2: What are new research trends regarding the research topic of this paper?

The rest of the paper includes as follows:

Section 2 presents related works and motivation; section 3 presents the studies' review to answer the earlier mentioned RQs, and section 4 concludes the paper.

2. Related Works and Motivation

Some articles about EHR, PHR, Blockchain, and security have been reviewed in this section that highlights our research requirements. However, we have also reviewed some of the papers relevant to our research topic for emphasizing our necessity to perform this study and to identify the research gap.

Researchers in [14] investigated the novelty of available privacy and security remedies for green IoT-based agriculture. They provided an overview of four-tier green IoT-based agriculture architecture. They could rank the menace models in contrast to green IoT-based agriculture into five groups, including authentication, attacks against privacy, availability, confidentiality, and integrity aspects via extensive analysis and research that had been performed. Besides, they analyzed the consensus algorithms and privacy-oriented blockchain-based remedies for green IoT-based agriculture. There yet numerous challenging areas have been available, like machine learning techniques, data sets for intrusion recognition, how to take the best consensus algorithm, scalability analysis of blockchain-based remedies, and the schematization of experimental and compatible cryptographic protocols. All these areas need to be checked out soon. Although their study has been systematic, it is only in agriculture; that requires to be performed in the research area of the current paper.

Ekramifard et al. in [15] conducted SLR on the blockchain innovation utilization in giving IoT security and protection. Because of the decentralized blockchain idea, its intrinsic secrecy was managed and gave network security to untrusted parties. Increasing extraordinary consideration in the difficulties of security concerns in IoT was identified. Their investigation showed that encryption that is utilized in blockchain-based methods is a workforce and time-consuming process. On the other hand, their finding showed that IoT gadgets have various capacities for processing altogether. Not every one of them is competent for running the encryption algorithms at the proper velocity. Scalability is a significant difficulty there because Blockchain includes a decentralized nature. Ledger size may elevate after some time, and as a rule, information size is larger than the most IoT nodes' limit.

Sultan and Mushtaq [16] investigated the literature on Blockchain and IoT and emphasized issues linked to an IoT atmosphere. Their analysis in the literature review showed that the IoT is the subsequent appearing technology and the enhancement of high-velocity networks and smart network devices. Unluckily, the tools of IoT are more disposed to attacks and not able to save

themselves. Their work highlighted various blockchain network aspects and features such as to eliminate the IoT problems. Besides, issues that are not addressed after blockchain execution are determined. Given that, their study was not done systematically. Therefore, this limitation has provided the context for this research.

3. Blockchain and Internet of Things

3.1. Blockchain Technology (RQ1)

Blockchain innovation is considered a moderately new creation; it rotates around utilizing an open changeless ledger named a blockchain that keeps records of a digital transaction [17-18]. Blockchain has been characterized as a peer-to-peer smart ledger of dealings that may be openly or secretly scattered to all members or clients in a decentralized structure. It utilizes cryptography procedure and consent to confirm dealings/exchanges that ensure the transaction authority keeps away from twofold spending and allows high-value transaction methods in a speculated mood [19]. Besides, it offers accuracy and cancels the requirement for third-parties/intermediaries directors [20]. Transaction accounts have been increasingly straightforward by blockchain innovation. This fact is because Blockchain is a kind of disseminated ledger where the total network members divide similar documentation and not various individual duplicates. The divided documentation adaptation can be refreshed distinctly through agreement.

3.2. Internet of Things (RQ1)

In the cutting-edge advanced condition, today, we are progressively associated with the web empowered electronic gadgets. The web has pushed toward getting omnipresent, impacting human life unbelievably by observing and controlling a wide assortment of hand-held gadgets. This fact creates enormous alteration in the manner of living things (objects) to non-living things (objects) by presenting highlights like cooperation with this present reality and intelligence to non-living creatures. They have been connected with the physical world and work cooperatively to achieve troublesome tasks that require a significant smartness level. In order to be in connection with the physical world, actuators and sensors have been utilized by the computerized world [21].

3.3. Security for the Internet of Things (RQ1)

The utilization of default or hardcoded passwords has been considered a significant issue with IoT security that can prompt security breaks. IoT gadgets frequently contain a limited resource and do not possess the necessary process resources to execute powerful security.

The linkage of the legacy asset not inalienably intended for the IoT network is another security issue. The substitution of the legacy asset with associated innovation may be cost-restrictive; so, vast numbers of benefits will be retrofitted with intelligent sensors. Also, the same amount of IoT gadgets remains in the system for a long time; the inclusion of security may be testing. An absence of industry-acknowledged guidelines additionally torments the security of IoT. There is a lack of single agreement upon the system, while numerous IoT security structures exist. Its combination and operational innovation systems have made numerous problems for security groups, particularly those entrusted with securing frameworks and guaranteeing end-to-end security in regions outside their domain. An expectation to absorb information is included, and IT groups with the best possible varieties of abilities ought to be placed responsible for IoT security. In **Fig. 1**, the applications of Blockchain in a different domain are shown.

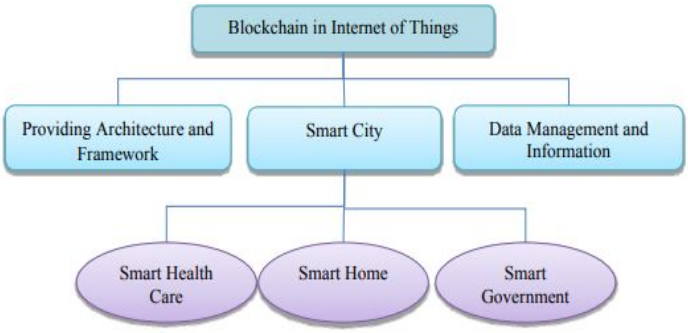


Fig. 1. Blockchain and Internet of Things

3.4. Security Challenge of Blockchain in the Internet of Things (RQ2)

The always-increasing and continually expanding numbers of IoT gadgets in the general public create a protected, open, and dependable infrastructure to preserve and execute the figured data. The introduced model for IoT makes utilization of a central cloud server model that results in single-point failures [22]. Evading the mentioned blockchain innovation is coordinated with IoT because Blockchain utilizes a distributed network. Blockchain innovation contains a huge advancement in the IoT security field [23]. The innovation of blockchain technology may be introduced from diverse perspectives and can be utilized in various settings, from money related situations to security conditions [23-24]. Blockchain innovation would give better answers to the issues looked at by the systems of IoT. **Fig. 2** explains the types of IoT and blockchain applications.

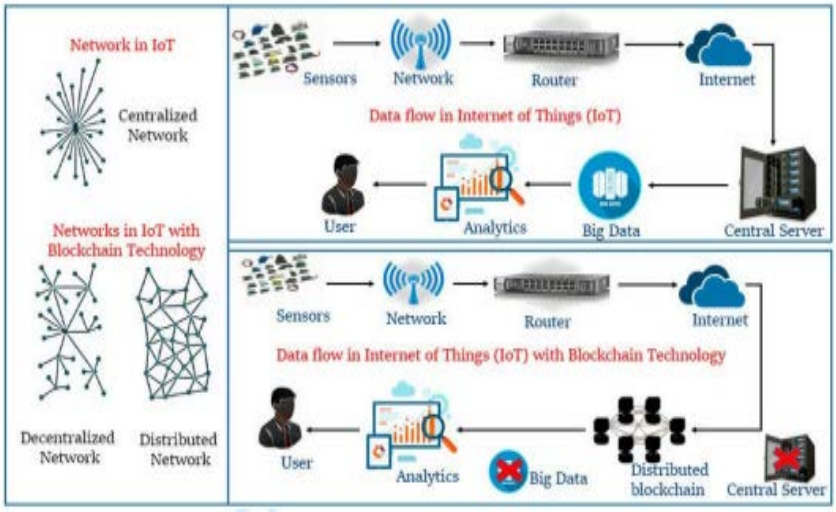


Fig. 2. IoT network types, data flow, and Blockchain

3.5 Research Trends and Challenges (RQ2)

Scalability: The examination of available and schematization of novel Blockchain for the scalable IoT systems are of great importance [25].

Lightweight Architectures and Schemes: Schematization and improvement of lightweight blockchain-based architectures for IoT systems are critical for decreasing the traditional blockchain overhead. Albeit, the identical privacy and security stage as customary Blockchain should be certified [26].

Computational Power: The systems of IoT containing an extended potency range are different. The conduction of encryption by the total nodes of IoT might not be feasible in experimental frameworks. So, other procedures must be offered for carrying encryption out with the utilization of an IoT nodes group or procedure that contains the least overhead on IoT nodes [27].

Storage: Blockchain's technology is proper for IoT systems that are decentralized since it does not contain a centralized controller. Albeit, all of the IoT nodes require saving the ledger that enhances in size with time. The nodes of IoT might not be able to save a huge data amount [28].

Optimal Design: An optimal IoT system must be schematized and take into account blockchain-depended privacy and security as a necessary factor. It may cause the creation of an optimal schematization that donates the same priority to computation, connectivity, privacy, coordination, and security [28].

Legal Issues: The standards of privacy and security differ in various regions. This fact is considered a crucial issue for the prosper blockchain technology conformity in the systems of IoT. There is a requirement for the standard scenario in which producers can make utilization of it to supply privacy and security remedies [29].

4. Conclusions

This study has concentrated on the technologies of Blockchain that are able to address numerous IoT security problems. The provided review emphasized the modern IoT security difficulties and weak-points in the Blockchain and IoT security systems. The taxonomy has been illustrated regarding the security of IoT and blockchain difficulties in the sight of supplied framework and architecture, smart city and data management, and information. Furthermore, the demonstrated facts are IoT security issues. The present investigation aims to explain the security challenges in blockchain technology in IoT systems. Ultimately, it points out the subsequent directions of blockchain technology and its effects on the IoT systems that may be useful for future developments.

References

[1] L. Wu, X. Du, W. Wang, and B. Lin, "An out-of-band authentication scheme for internet of things using blockchain technology," in *2018 International Conference on Computing, Networking and Communications (ICNC)*, 2018: IEEE, pp. 769-773.

[2] T. Alam, "IoT-Fog: A communication framework using blockchain in the internet of things," *arXiv preprint arXiv:1904.00226*, 2019.

[3] H.-N. Dai, Z. Zheng, and Y. Zhang, "Blockchain for Internet of Things: A survey," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8076-8094, 2019.

[4] D. E. Kouicem, A. Bouabdallah, and H. Lakhlef, "Internet of things security: A top-down survey," *Computer Networks*, vol. 141, pp. 199-221, 2018.

[5] Y. Yu, Y. Li, J. Tian, and J. Liu, "Blockchain-based solutions to security and privacy issues in the Internet of Things," *IEEE Wireless Communications*, vol. 25, no. 6, pp. 12-18, 2018.

[6] K. Biswas and V. Muthukkumarasamy, "Securing smart cities using blockchain technology," in *2016 IEEE 18th international conference on high performance computing and communications; IEEE 14th international conference on smart city; IEEE 2nd international conference on data science and systems (HPCC/SmartCity/DSS)*, 2016: IEEE, pp. 1392-1393.

[7] H. Damghani, L. Damghani, H. Hosseinian, and R. Sharifi, "Classification of Attacks on IoT," in *4th International Conference on Combinatorics, Cryptography, Computer Science and Computation*, 2019.

[8] H. A. Abdul-Ghani, D. Konstantas, and M. Mahyoub, "A comprehensive IoT attacks survey based on a building-blocked reference model," *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 9, no. 3, pp. 355-373, 2018.

[9] P. Williams, P. Rojas, and M. Bayoumi, "Security Taxonomy in IoT–A Survey," in *2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2019: IEEE, pp. 560-565.

[10] T.-H. Nguyen and M. Yoo, "A hybrid prevention method for eavesdropping attack by link spoofing in software-defined Internet of Things controllers," *International Journal of Distributed Sensor Networks*, vol. 13, no. 11, p. 1550147717739157, 2017.

[11] Y. Lu and L. Da Xu, "Internet of things (iot) cybersecurity research: A review of current research topics," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2103-2115, 2018.

[12] G. Rajendran, R. R. Nivash, P. P. Parthy, and S. Balamurugan, "Modern security threats in the Internet of Things (IoT): Attacks and Countermeasures," in *2019 International Carnahan Conference on Security Technology (ICCST)*, 2019: IEEE, pp. 1-6.

[13] N. Kshetri, "Can blockchain strengthen the internet of things?," *IT professional*, vol. 19, no. 4, pp. 68-72, 2017.

[14] J.-P. A. Yaacoub *et al.*, "Securing internet of medical things systems: limitations, issues and recommendations," *Future Generation Computer Systems*, vol. 105, pp. 581-606, 2020.

[15] A. Ekramifard, H. Amintoosi, A. H. Seno, A. Dehghantanha, and R. M. Parizi, "A systematic literature review of integration of blockchain and artificial intelligence," in *Blockchain Cybersecurity, Trust and Privacy*: Springer, 2020, pp. 147-160.

[16] A. Sultan, M. A. Mushtaq, and M. Abubakar, "IOT security issues via blockchain: a review paper," in *Proceedings of the 2019 International Conference on Blockchain Technology*, 2019, pp. 60-65.

[17] F.-J. Ferrández-Pastor, H. Mora, A. Jimeno-Morenilla, and B. Volckaert, "Deployment of IoT edge and fog computing technologies to develop smart building services," *Sustainability*, vol. 10, no. 11, p. 3832, 2018.

[18] C. C. Agbo, Q. H. Mahmoud, and J. M. Eklund, "Blockchain technology in healthcare: a systematic review," in *Healthcare*, 2019, vol. 7, no. 2: Multidisciplinary Digital Publishing Institute, p. 56. [19, 20]

[19] T. C. Stratopoulos, V. X. Wang, and J. Ye, "Blockchain technology adoption," *Available at SSRN 3188470*, 2020.

[20] T. L. Lee, T. L. Lim, B. Y. Teh, J. P. Tou, and N. L. Wong, "Cryptocurrency Bitcoin: fundamental drivers explained," UTAR, 2018.

[21] M. Banerjee, J. Lee, and K.-K. R. Choo, "A blockchain future for internet of things security: a position paper," *Digital Communications and Networks*, vol. 4, no. 3, pp. 149-160, 2018.

[22] S. Rathore, B. W. Kwon, and J. H. Park, "BlockSecIoTNet: Blockchain-based decentralized security architecture for IoT network," *Journal of Network and Computer Applications*, vol. 143, pp. 167-177, 2019.

[23] J. Zhou, Z. Cao, X. Dong, and A. V. Vasilakos, "Security and privacy for cloud-based IoT: Challenges," *IEEE Communications Magazine*, vol. 55, no. 1, pp. 26-33, 2017.

[24] O. Novo, "Blockchain meets IoT: An architecture for scalable access management in IoT," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1184-1195, 2018.

[25] W. Ejaz and A. Anpalagan, "Blockchain technology for security and privacy in Internet of things," in *Internet of Things for Smart Cities*: Springer, 2019, pp. 47-55.

[26] L. Zhou, X. Li, K.-H. Yeh, C. Su, and W. Chiu, "Lightweight IoT-based authentication scheme in cloud computing circumstance," *Future Generation Computer Systems*, vol. 91, pp. 244-251, 2019.

[27] Y. Ren *et al.*, "Data query mechanism based on hash computing power of blockchain in Internet of Things," *Sensors*, vol. 20, no. 1, p. 207, 2020.

[28] M. U. Hassan, M. H. Rehmani, and J. Chen, "Privacy preservation in blockchain based IoT systems: Integration issues, prospects, challenges, and future research directions," *Future Generation Computer Systems*, vol. 97, pp. 512-529, 2019.

[29] J. Lindman, V. K. Tuunainen, and M. Rossi, "Opportunities and risks of Blockchain Technologies–a research agenda," 2017.

Positive Experience Design Methods of IOT Intelligent Product

Chun-mao Wu, Hua-yuan Xu
Department of Product Design, Donghua University
Shanghai, China
[e-mail : cmwu@dhu.edu.cn]
*Corresponding author: Chun-mao Wu

Abstract

Objective This paper proposes a positive experience design method for IoT intelligent products to improve users' subjective well-being in the background of artificial intelligent and big data. **Methods** Firstly, selecting six target users and taking the MI IoT intelligent products as the research objects to do deep observation about the using IoT intelligent products in their own home in 2 weeks by the methods of home-visiting interview, group diary and focus group. Secondly, constructing the individual activities table of participants' IoT intelligent product experience by Hierarchical Task Analysis (HTA). Thirdly, two design scholars classify and extract the sub-tasks of happiness in HTA table. Finally, getting the positive experience design method of IoT intelligent products. **Conclusion** The positive experience design method of IoT intelligent products is proposed from focusing on personal pleasure experience to individual life meaningful design and group social relationship design, including individual pleasure experience, personal goal realization, group needs satisfaction and the harmony of group relations. In the era of big data, it is helpful for designers to use the design method to improve users' sense of sustainable pleasure, achievement perception of future goal realization and happiness of group social relationship.

***Keywords*:** Positive experience; IoT Product; Artificial Intelligent; Design Method

1. Introduction

In recent years, the Internet of things has been involved in people's daily life, work and study. Through the use of products, networks and sensors, powerful applications and services have been produced, and people's lifestyle has been changed. However, the research on the design of Internet of things (IOT) products at home is still little [1] . The research on how IOT product design can improve people's subjective well-being is still insufficient. The purpose of this paper is to build a positive experience design method of Internet of things (IOT) products which is conducive to the improvement of subjective well-being, so as to help designers' experience design practice more effectively from Internet of things products under the guidance of the method.

2. Literature Review

2.1 IoT Product

The concept of IoT can be traced back to Bill Gates' book the road Ahead, that published in 1995; in 1999, Kevin Ashton, co-founder of MIT's automatic identification center, formally proposed the concept of IoT, that is, the information of all objects is connected to the Internet through radio frequency identification

This research was supported by a research grant from Shanghai Pujiang Program (No. 19PJC003)

(RFID) sensor devices for intelligent management and identification; in 2005, the International Telecommunication Union (ITU) pointed out that the ubiquitous "Internet of things" Communication Era is coming, and all the goods in the world, from tires to toothbrushes, from houses to tissues, can be connected for interaction[2]. In short, the Internet of things is a new technology based on the Internet. It combines goods with the Internet to form a huge network to realize the interconnection of people, machines and things at any time and place.

IoT products are composed of sensor equipment, communication infrastructure, cloud computing, processing unit, decision-making and action call system. These components and systems cooperate with each other through their own unique functions to realize the highly intelligent and automatic products. Users can improve the quality of life according to the information provided by the IoT products [3]. At present, IoT products have been widely used in various aspects: in the field of automobile, sensors are installed to provide drivers and maintenance personnel with relevant potential risk information [4]. In the field of smart home, the air conditioner and purifier in the home can be remotely controlled through the Internet, so as to create a more convenient, safe and comfortable home living environment [5]. In the medical field, intelligent wearable is adopted Wear the sensor in the device to sense the patient's physical condition and contact the hospital for diagnosis when necessary [6].

2.2 Positive Experience Design

Positive experience originates from positive psychology, which is a discipline that studies the optimal operating conditions and processes of people, groups and institutions, and contributes to human flourishing [7]. Positive psychology focuses on what makes life valuable and what conditions determine human well-being. It studies the things that people think are meaningful in life, even the factors that make people happy, so as to solve the problems in human life [8]. Positive experience design is based on positive psychology, through the positive design of intervention behavior, in the process of interaction between people and products or services, give people a happy and positive experience, this experience is not only conducive to individual development, but also conducive to the flourishing of the environment or community. It is not only satisfied with short-term happiness, but also has a long-term positive impact on individual development [9].

The purpose of positive experience design is to improve people's subjective well-being, which can be realized through two dimensions of pleasure and meaningful design. Hassenzahl [10] proposed six basic psychological needs as the source of positive product experience; Wiese [11] pointed out that sustainable happiness depends more on user behavior than on material wealth. Wu Chun-mao [12] put forward the concept design canvas based on positive experience, which helps designers to use the tool to capture users' positive design opportunities in a standardized and rapid manner. From the perspective of meaning design, this paper studies the significance of design to improve the subjective well-being of users. Orth [13] provides an application of product attachment theory in the design process of customized products, and verifies the importance of forming meaningful association between objects and individuals through design case study; Casais [14] based on the symbolic meaning of products can promote people's happiness, and puts forward 16 design directions that can be used for conceptualization; Ozkaramanli [15] proposes three design strategies of improving subjective wellbeing driven by the self-control dilemma contradiction of short-term pleasure and long-term wellbeing. Based on the above literature research, this paper proposes a positive experience design method of IoT products to improve subjective well-being through the two-week user experience research of IoT products.

3. Research Method

3.1 Target Product

In this study, Xiaomi IoT products are selected as the research objectives (Fig. 1), because it not only conforms to the basic characteristics of IOT products, but also is a typical representative of China's smart home, and has a wide range of user groups. It is convenient for the author to recruit target users and observe how users perceive, experience and interact with IoT products, so as to find out the motivation behind the interaction between users

Fig. 1. Xiaomi Internet of Things Products

and IoT, and the interaction factors that affect users, so as to contribute knowledge for future IoT product design practice.

Xiaomi smart home system, with Xiaoai speaker as the control center, uses the Internet of things technology to connect daily home products. Through the voice interaction with Xiaoai speaker, the intelligent home products such as lighting equipment, air conditioner and air purifier can be controlled, and the control instructions can be customized according to the application scenarios. In addition, Xiaomi smart home series also includes a variety of sensors to achieve highly intelligent products. For example: air conditioning is controlled by temperature and humidity sensor, light is controlled by human body sensor, and air purifier is controlled by door and window sensor. After the initial setting of the above products, there is no need for user operation, and the products will coordinate with each other to meet the needs of users.

3.2 Participants

This study recruited six participants from different families and using Xiaomi IoT products in the local community. These six participants are not expert users of IoT products. As shown in Table1, the age, occupation, living conditions and familiarity with IoT products of the six participants were different. The participants were divided into two groups: group 1 (G1) with high familiarity with digital products and group 2 (G2) with low familiarity with digital products. All the research process was completed in the participant's home through two weeks of in-depth observation. Meanwhile, the behavior and thinking of participants were recorded by video and questionnaire.

Table 1. Participant Profiles

	NO.	Profile (gender, age, occupation)	Lives with	Knowledge of the IoT	Digital familiarity
G1	P1	Female,26, student	roommate	Heard of it	High
	P2	Male,25, student	Parents	Xiaomi Products	High
	P3	Female,35, office worker	Husband, Child and parents	None	High
	P4	Female,26, office worker	Husband	Parts of Xiaomi products	High
G2	P5	Male,42, office worker	Wife and 2 kids	None	Low
	P6	Female,45, housewife	Husband and Child	None	low

3.3 User Research Methods

In order to understand how the participants, perceive and experience the IoT products, this study uses the methods of home interview, group diary and focus group, as shown in Figure 2.

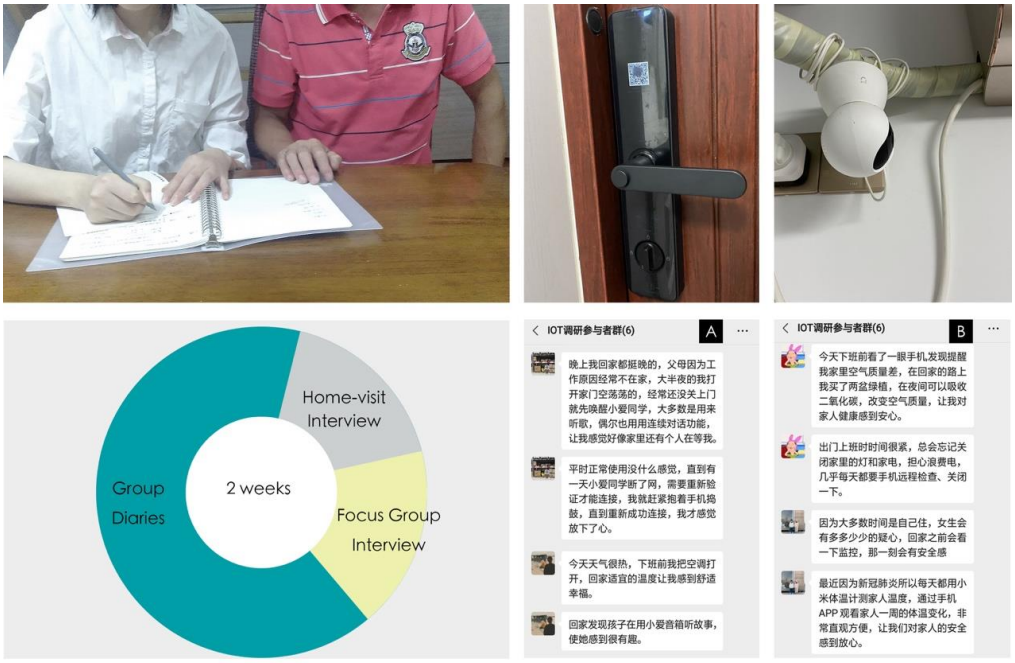
First of all, through the household interview, the author understands the perception and expectation of the participants. A semi-structured family interview was conducted the day before the event to get a general understanding of the use of IoT products.

Secondly, through the group diary, record the user's experience and perception in the using process. Participants are required to share their thoughts and feelings related to IOT products through social media (we-chat, QQ) at least every two days, and other participants can leave messages and communicate. The purpose is to help participants to have the opportunity to think about functions they didn't expect in use, and hope that the posts can stimulate feedback from participants.

Third, through focus groups, participants share their past experiences and get more details behind the experience based on happy moments.

Fourthly, based on video, family interview and social media sharing, the author uses the hierarchical task analysis (HTA) method to get the participants' personal activity table, which provides an objective and typical description of the relevant sub units of the Internet of things product using activities.

Fifthly, according to the subjective statement of the source of well-being in the HTA summary table, two design researchers further classified and extracted the sub task column with happiness, and then obtained the positive experience design method.



Translated:

A . 1. I go home late in the evening. My parents are often not at home because of work. In the middle of the night, I open the door and wake up Xiaoai before I close the door. Most of times are used to listen to songs and occasionally use the dialogue function, which makes me feel like there is someone waiting for me at home.
2. Usually I could not feel the value of Xiaoai, until one day Xiaoai is cut off the network, and need to re verify to connect. I quickly hold the mobile phone to deal with, until the successful connection, I feel at ease.
3. It's very hot today. I turn on the air conditioner before arriving at home. The suitable temperature makes me feel comfortable and happy.
4. I find my kid is listening the story with Xiaoai sound box, and make her happy.

B . 1. Today, I took a look at my mobile phone before arrived at home and found that the air quality was poor. I bought two pots of green plants, which could absorb carbon dioxide at night, and changed the air quality, and made me feel at ease about the health of my family.
2. When I go out to work, I always forget to turn off the lights and appliances at home. I have to check my mobile phone remotely almost every day.
3. Because most of the time I live lonely in my apartment, and worry about my safety. I always have a look at the monitoring before arriving at home, give me a sense of security at that moment.
4. Recently, because of COVID-19, the family members' temperatures are measured by XIAOMI thermometer every day. The temperature changes chart in a week can be checked by APP, give me a sense of security.

Fig. 2. Study process diagram

4. Result Analysis

At the end of the study, the author collected the group diaries from social media, including photos and communication records, as well as focus group conversation records. The data was sorted out and analyzed by analytic hierarchy process. Finally, six hierarchical task analysis forms were obtained, corresponding to the Internet of things product experience of the six participants. Table 2 shows the HTA table for participant 3. Based on HTA and semantic analysis, the author will introduce the personal IoT product experience results of six participants.

Participant 1's IoT product experience process is composed of 9 task elements, which are divided into 24 subtasks, and the number of subtasks corresponding to each task is different. Among all the sub tasks, 6 were marked as the sub tasks to obtain happiness, namely, voice interactive alarm clock, Xiaoai music, voice interactive learning cooking, viewing the working screen of sweeping robot by mobile phone, learning and reminding air quality. For example, when she gets up in the morning and uses the alarm clock function of Xiaoai speaker, she thinks it is very interesting and pleasant to control the alarm clock to stop by voice. She thinks that getting up on time will also make her feel autonomy. She thinks that this is an important step to continuously realize her personal goals.

Participant 2 claims to be a fan of intelligent products and pays close attention to the development of current scientific and technological products. He recorded a total of 7 IoT product using tasks, which were divided into 22 subtasks. In the Rubik's cube controller experience task, he summarized the three subtasks as one with a sense of happiness (rotating the Rubik's cube to control lamp brightness, knocking twice the curtain to open automatically, flipping to let all the home appliances to turn off automatically), which was recorded as interactive operation. He thinks that it is convenient and interesting to control household products by rotating, shaking and knocking the Rubik's cube controller. In addition, participant installed a variety of sensors in their homes, forming a powerful IoT. Most products have subtasks that make him feel happy.

Participant 3 lived with her children and parents, and the study found that most of her sources of happiness were related to her families. She filled in six tasks, which were divided into 25 subtasks, of which 7 were marked as having happiness, namely, the sensor automatically turned on the light, the dishwasher checked the water consumption data, took the dishes, the purifier worked, looked at the purifier feedback data, Xiaoai interactive storytelling, and mobile phone viewing the monitoring screen. For example, when using the air purifier, you will check the air quality through the mobile app every day. When the air quality index drops suddenly, you will find out the reason or take some recovery measures. If you see good data, you will feel at ease about your family's health. In addition, through the storytelling interaction of Xiaoai speaker with children and family to increase children's interest in learning; through the mobile phone to check and monitor when you are at work, at any time care about the status of the elderly at home, voice and video interaction, these are the happiness time for P3 to create a harmonious atmosphere for their families.

During the interview, the author learned that participant 4' families are very concerned about sustainable development and environmental protection. They usually pay attention to the electricity, water quantity and garbage classification of their homes. Among the six sources of happiness (mobile phone remote switches off, intelligent lock closed, home automatic power cut-off, Xiaoai reminding to clean up, voice TV frequency modulation, mobile phone sliding control light brightness, inspection and supervision Control screen) can also be seen. For example, when you go to work, you can turn off the lights and air conditioner through your mobile phone that you forgot to turn off at home; the door lock of the IoT opens automatically, and when you close the door, light will automatically be turned off. She believes that these functions can not only save family expenses, but also save resources, and reflect the sense of social responsibility.

Table 2. Overview of P3's HTA

HTA focus		Subjective Description	
Task（number of subtasks）	Subtask	Subtask with Well-being	Source of Well-being
Somatosensor（1）	1.Turn on the light	1.Turn on the light	1. When elder goes to the toilet in the midnight, the sensor of light feels that the human opens automatically, which makes me feel safe and at ease.
Wash dishes（6）	1. Put the dishwashing powder; 2. Put in the bowl; 3. Press the power; 4. Waiting; 5. The mobile phone reminds you to finish; 6. Take out the bowl;	1.The mobile phone reminds you to finish;	1. After washing dishes, the mobile phone app alerts and displays the water consumption, time and so on, which increases the trust to the machine.
		2.Take out the bowl;	2. After sterilization, when taking out the bowl, it still feels warm, and there is a sense of happiness that the health of the family is guaranteed.
Air purification（6）	1. Plug in the power; 2. Xiaoai starts the purifier; 3. Set the mode; 4. Enjoy the air purification; 5. check the data by the mobile phone; 6. Close the purifier;	1. Enjoy the air purification; 2. check the data by the mobile phone;	1. When the purifier works, you can inquire about the current air quality through Xiaoai at any time. 2. The accurate data from the air purifier reassured me about the health of my family.
Sweep（4）	1. Timing the sweeping robot; 2. Checking cleaning information by mobile phone; 3. Reminding completion; 4. Automatic charging;		
Xiaoai speaker（5）	1. Wake up; 2. Choosing story by voice; 3. Children listen to story 4. Elders listen to opera; 5. Voice stop;	1.Choosing story by voice	1.Children choose the story they like by voice interaction; this process makes kids happy.
View monitoring（3）	1. Turn on the mobile phone; 2. View the monitoring screen; 3. Turn off the mobile phone	1.View the monitoring screen	1.When you go to work, you can check the status of the children and the elderly at home by voice interaction, which is more reassuring.

Participants 5 and 6 are both low digital people. In the interview, they don't know what the IoT is before using IOT products, so they have low expectations and trust in the IoT. From their HTA forms, it is found that the data feedback by IOT products will give them a sense of reliability. For example, when using the rice cooker, the mobile phone software can check the working process at any time; when using the air purifier, it can view the air quality data at home; after the dishwasher completes its work, it will feed back the data such as water consumption and time to users. These data feedbacks promote product reliability and improve the sense of security and trust of people with low digital level.

Fig. 3. Sub-tasks classification of Well-being

The happiness sources of participants 5 and 6 also included Xiaoai's "welcome home" voice, automatic light adjustment mode, family members watching movies together, desk lamp rotating button to control brightness, pushing mopping machine to mop the floor, remotpower failure of mobile phone, and selection of programs by Xiaoai speaker. For example, P5 go home from work every day. After opening the door through the intelligent lock, Xiaoai speaker will automatically play the voice of "welcome home", which gives him the feeling of accompany. In the view of P6, compared with the sweeping robot, the use of human-computer interaction makes her feel more confident about the cleanliness of cleaning. It can not only easily complete the cleaning, but also take the practice, and feel happy.

Based on the above data results, two design researchers further classified the sub tasks with happiness into four categories according to the source of happiness, which were defined as personal pleasure experience, individual goal realization, group demand satisfaction and collective relationship harmony (see Figure 3).

5. Discussion

Based on the above research on user experience of IoT products, combined with previous literature research, the author further sums up the positive experience design method of IoT products from two dimensions of design object (individual, family) and source of happiness (pleasure and meaning). The four design directions are further summarized into the positive experience design methods of IoT products shown in Figure 4. Design objects: in the design of IoT products, the design faces not only a single individual, but also multiple individuals with family as the unit, and there are complex social interactions between individuals. Therefore, designers should pay attention to two design objects (individual and family) in the positive experience design of IoT products. The source of happiness: designers can get inspiration from two sources of happiness: pleasure and meaning. Pleasure refers to the short-term satisfaction of needs through happy and comfortable experience; meaning refers to the subjective well-being obtained through the purposeful pursuit of self-realization or life value. The details are as follows.

(1) Individual pleasure experience. To achieve the purpose of enjoying happy experience through positive design, that is, individuals obtain subjective well-being through short-term happiness. The implementation method can be achieved by improving the disadvantages of existing products, or by strengthening the existing sources of happiness or introducing new sources of happiness. For example, from sweeping robot can only sweep the floor to realize the function of sweeping and dragging, the convenience brought to us will make people happy; through the voice control of home products by Xiaoai speaker, this new interactive way will bring us fun and pleasure.

(2) Personal goal realization. To obtain subjective well-being through the realization of long-term goals with personal significance. Design is no longer concerned with short-term happiness, but to achieve the goals and aspirations in a long-term. The source of subjective well-being can be the sense of achievement of personal goals, or the sense of

progress towards future goals. For example, through the interactive voice search of learning materials with Xiaoai speaker, you will get a sense of wellbeing if you achieve excellent results after a period of time; you will feel happy if you get up and Study on time through Xiaoai's alarm clock every day and constantly strive for your own learning goals.

(3) Group needs satisfaction. Based on the current situation that the design object of IoT for family is multiple individuals, the design should pay more attention to the significance of group. Tsai et al.[16] proposed that the future IoT product design should be balanced between family members' conflicting and personal values. Positive experience design of IoT products should be designed to meet the experience needs of all members of the family, so as to achieve subjective well-being. For example, in the family of P3, Xiaoai speaker has become a very popular product. She usually uses it to control the intelligent products at home, while her mother uses it to listen to operas. She often accompanies her children to listen to stories and have some interesting dialogues and interactions. Xiaoai meets the different experience needs of the whole family members.

(4) **Harmonious group relationship.** Creating a good family atmosphere through positive experience design, that is to mobilize the positive communication between family members through an IoT product design, to promote the harmony of family members and obtain subjective well-being. For example, in this study, P5 and P6 often organize family members to watch movies together on home TV, and provide conditions for everyone to discuss, so as to make it a platform carrier for building a harmonious group relationship.

(5) Each of the above directions can affect the positive experience design results of IoT products and stimulate the generation of subjective well-being of users. What needs to be pointed out is that positive experience design is not only limited to the realization of design objectives in one direction, but also can meet multiple directions at the same time. For example, through the learning function of Xiaoai speaker, it not only gives users a pleasant experience, but also achieves personal goals. Designers can choose the design methods according to different design situations, which helps to inspire IoT product designers to bring users continuous pleasure and enhance their subjective well-being.

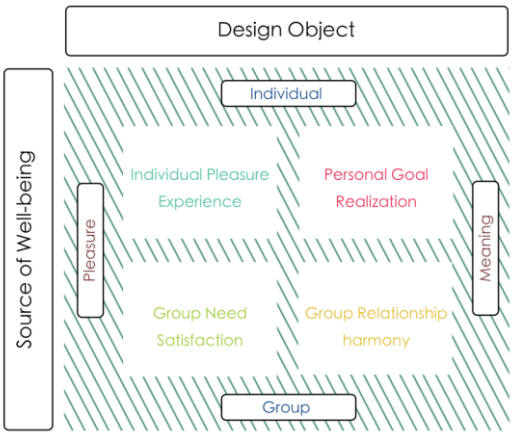


Fig. 4. Positive Experience Design Method of IoT Product

6. Conclusions

The contributions of this paper are shown as follows. Based on the relevant theory of positive experience design, through the two-week study on the use of Internet of things products of users, this paper proposes a positive experience design method of IoT products to improve the subjective well-being of users, including four IoT product design directions: personal pleasure experience, individual goal realization, group demand satisfaction, and harmonious group relationship. On the basis of the research method, subsequent designers can carry out corresponding IoT product design practice to verify the feasibility and effectiveness of the design method, and contribute the relevant positive experience design knowledge.

References

[1] CHO M, LEE S, LEE K P, "How do People Adapt to use of an IoT Air Purifier?: From low Expectation to Minimal Use," *International Journal of Design*, Vol13, No. 3, pp.21-38,2019.

[2] SAYAR D, ER O, "The Antecedents of Successful IoT Service and System Design: Cases from the Manufacturing Industry," *International Journal of Design*, Vol 12, No. 1, pp. 67-78, 2018.

[3] GHAJARGAR M, WIBERG M, STOLTERMAN E, "Designing IoT Systems that Support Reflective Thinking: A Relational Approach," *International Journal of Design*, Vol 12, No. 1, pp. 21-35, 2018.

[4] JAE J C, HYUN J K, "An Automobile. Environment Detection System Based on Deep Neural Network and its Implementation Using IoT-Enabled In-Vehicle Air Quality Sensors," *Sustainability*, Vol 12, No. 6, pp. 2475-2492, 2020.

[5] FENG S, SETOODEH P, HAYKIN S, "Smart Home: Cognitive Interactive People-Centric Internet of Things," *IEEE Communications Magazine*, Vol 55, No. 2, pp. 34-39, 2017.

[6] DWIVEDI A D, SRIVASTAVA G, DHAR. S, et al, "A Decentralized Privacy-Preserving Healthcare Blockchain for IoT," *Sensors*, Vol 19, No. 2, pp. 326-343, 2019.

[7] SNYDER C R, LOPEZ S J, "*Handbook of Positive Psychology*," New York: Oxford University Press, 2002.

[8] GABLE S L, HAIDT J, "What (and Why) is. Positive Psychology?" *Review of General Psychology*, Vol 9, No. 2, pp. 103-110, 2005.

[9] WU Chun-mao, WEI Wei, Li Pei, "Positive Design Model for Improving Subjective Well-being," *Packaging Engineering*, Vol 40, No. 12, pp. 29—33, 2019.

[10] HASSENZAHL M, ECKOLDT K, DIEFENBACH S, et al, "Designing Moments of Meaning and Pleasure: Experience Design and Happiness," *International Journal of Design*, Vol 7, No. 3, pp. 21-31, 2013.

[11] WIESE L, POHLMEYER A, HEKKERT P, "Activities as a Gateway to Sustained Subjective Well-Being Mediated by Products," *Proceedings of the 2019 on Designing Interactive Systems Conference*, 2019.

[12] WU Chun-mao, ZHANG Xiao-nan, WU Xiang, "Construction of Conceptual Design Canvas Based on Positive Experience," *Packaging Engineering*, Vol 41, No. 16, pp. 76—82, 2020.

[13] ORTH D, THURGOOD C, HOVEN E, "Designing Objects with Meaningful Associations," *International Journal of Design*, Vol. 12, No. 2, pp. 91-104., 2018.

[14] CASAIS M, MUGGE R, DESMET P, "Objects with Symbolic Meaning: 16 Directions to Inspire Design for Well-Being," *Journal of Design Research*, Vol 1, No. (3/4, pp. 247-281, 2018.

[15] OZKARAMANLI D, OZCAN E, DESMET P M A, "Long-Term Goals or Immediate Desires? Introducing a Toolset for Designing with Self-Control Dilemmas," *The Design Journal*, Vol 20, No.2, pp. 219-238, 2017.

[16] TSAI W C, CHUANG Y L, CHEN L L, "Balancing Between Conflicting Values for Designing Subjective Well-Being for the Digital Home," *Proceedings of APCHIUX*, pp. 33-36, 2015.

Research on Multi-channel Face Recognition Based on Siamese Network

Xibin Jia, Ntazana Mazimba Windi Siluyele, Meng Yukun and Yao Jigang
Beijing University of
Technology, Beijing, 100124, China
[e-mail: jiaxibin@bjut.edu.cn]
*Corresponding author: Xibin Jia

Abstract

With numerous security cameras present in airports, offices, universities, ATMs, Banks, and in other locations with a security system, facial recognition becomes one of the most suitable way to detect, analyze and process different people identities. However, most face recognition networks directly use full face region features, and local region features and this may hide more critical information. In this paper, Siamese network recognition algorithm is used to ensure the robustness of the algorithm process in detecting the different facial features. The local region features and global region features of human face are then extracted to be processed. For different local region features and global region features, adaptive weighting method is used to improve the recognition accuracy. The comprehensive experiments are done via Cassia-FaceV5 face database with the network we proposed
resulting in an accuracy of more than 90%.

Keywords: Facial recognition, Siamese network, Facial detection, biometrics, Face images.

1. Introduction

With the development of projects such as big data search, artificial intelligence, smart cities, and the use of corresponding technologies, our demand for face recognition technology for security protection is also increasing. Face recognition technology can be used as an important technical means to provide significantly more help to the relevant judicial departments in terms of capturing criminals and verifying identity documents. In addition, face recognition can also be used as a way to control access, such as the practical application of access control, time and attendance system, authorized entry, etc., to ensure that the identity information of entering and exiting personnel is recorded. At present, in human recognition, neural networks are mainly used for deep learning to realize recognition. In essence, neural networks are a method of learning based on data samples. After the use of neural networks for face recognition, human recognition technology has achieved great success. Shang-Hung Lin et al. trained three Probabilistic Decision Based Neural Networks (PDBNN) to detect faces and locate eyes for face recognition [1], finally completed a complete face recognition system.Yaniv Taigman et al. used Convolutional Neural Network (CNN) [2] and 3D modeling method based on reference points to maximize the effect of the model. In this paper, Siamese network is used to reduce the dimensionality of the face image by using the similarity metric (Similarity Metric Discriminatively) algorithm [3] to realize the siamese network structure for training face data. In order to strengthen the robustness, this paper also uses some local features as an auxiliary strategy for face identification, that is, first extracts the global features and local features of a person. In the above matching case, weighted evaluation is performed according to the credibility of model training, and in this form, local features are assisted to modify the global features, so as to achieve the reliability of face recognition under certain occlusion conditions.

2. Methods

2.1 Overview of multi-channel joint feature learning model

Since the face recognition part implemented in this paper is a single-sample recognition, it is also required to strengthen the robustness of the face recognition algorithm as much as possible on this basis,In order to achieve this goal,we propose a method of global and local feature joint verification as shown in **Fig. 1**. That is, other local features are used for recognition on the overall face recognition result, and the result is weighted with the overall recognition result. At the same time, when the local feature is not found in the local recognition, the weight of the part is reasonably reduced to reflect the part Robustness of face recognition algorithm in the case of location occlusion.

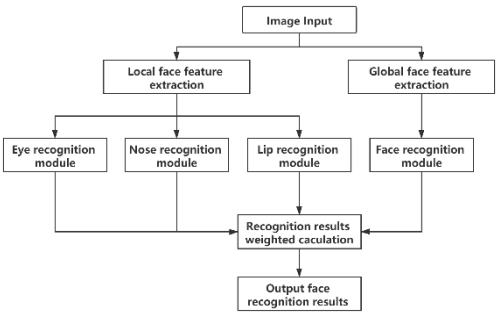


Fig. 1. The process of multi-channel face identification

2.2 Siamese network

Siamese network, in simple terms, is a connected neural network with two networks sharing weights. The network structure is shown in **Fig. 2**. Sharing weights means that the weights of the left and right neural networks are exactly the same, or even just one network without the need to implement another one. The function of the Siamese neural network is to measure the similarity of two images. These two neural networks respectively map the input to a new space to form a representation of the input in the new space. Through the calculation of the loss function, the similarity of the two inputs is evaluated. Let $Input_1$ and $Input_2$ represent a pair of images. Suppose that the output Y of the network is the binary label of the pair of images. If the images $Input_1$ and $Input_2$ belong to the same person (a "real pair"), then Y=0, otherwise Y=1 (an imposter). Let W be the shared weight to be learned, and let $G_w(Input_1)$ and $G_w(Input_2)$ be the two points generated by the mappings $Input_1$ and $Input_2$ in the low-dimensional space, then the loss function can be Expressed as

$$E_w(Input_1, Input_2) = \|G_w(Input_1) - G_w(Input_2)\| \quad (2.1)$$

Then $E_w(Input_1, Input_2)$ can represent the similarity between $Input_1$ and $Input_2$.

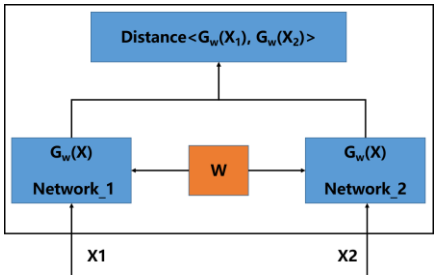


Fig. 2. The structure of siamese network $G_w(X_1)$

2.3 Haar features extraction

The idea of Haar classifier is to slide a sub-window in the image and calculate the feature value of the sub-window area, and then use the feature area of the trained classifier to filter, and the area that passes all the strong classifiers is the face. Haar-like features are divided into three types: edge features, line features and center-surround features as shown in **Fig. 3**. Using the haar-like features to train the face area and the non-face area. The white area minus the black area pixel value is the eigenvalue. We use a variety of feature matrices to train the classifier, face area and non-human area. The greater the difference in the feature values of the face region, the more appropriate the feature matrix is. The combination of the feature matrix, that is, the selection of the classifier is based on the AdaBoost algorithm.

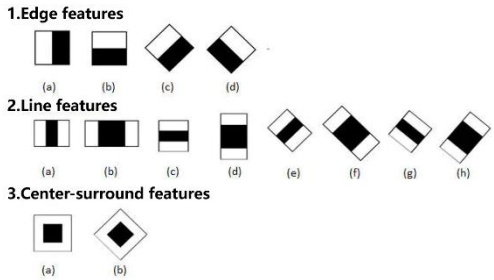


Fig. 3. Haar-like features

3. Experiments

3.1 Dataset

The main purpose of the application environment in this paper is to improve the robustness of the face recognition algorithm. This requires that the database has various face data images with different expressions, different illumination conditions and different occlusion conditions (such as wearing a hat, wearing glasses, etc.), on this basis, the CASIA-FaceV5 Asian face database is selected, which contains 2500 face images of 500 people , 5 different face images per person.

Positive and negative sample equalization: since the face recognition in this paper is a multi-feature joint recognition method, the overall face recognition is mainly used, and the local features (eyes, nose and mouth) are supplemented. The method of correcting the recognition results is based on the above process , Based on the storage of the global and local facial features, the path of each local feature data picture is constructed, so that it is convenient to read the positive and negative sample pairs needed in the training of the picture construction when using Siamese network. When constructing the path file of positive and negative sample pairs, this paper uses a dictionary tree to limit the number of negative sample pairs of the same person. The path that reads the image first is used as the label according to the person number in the path to limit the number of output negative sample pairs, so that the number of positive and negative pairs in the sample pair path file is close to 1:1.

3.2 Experimental Setup

In the experiment process, this paper uses a Siamese network constructed by a fully connected layer and a Siamese network constructed by a convolutional network using a convolutional layer and a pooling layer respectively, and compares the effects of the model before and after using positive and negative sample equalization. Some local features of the face are used, and a weighted strategy is adopted. The size of the weight is selected based on the accuracy of the model as the confidence space, and the weight of each part is set according to the experience value to modify the results of whole face recognition.

3.3 Result

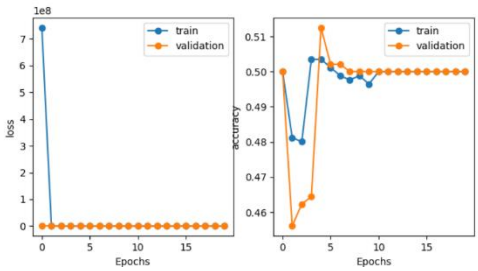


Fig. 4. Using full connection layer training

In the case of matching input with positive and negative sample pairs, that is, the ratio of positive and negative samples is 1:1. After many tests, it is found that the face image in the fully connected network cannot be trained normally, as shown in **Fig. 4**. After adding a convolutional layer and a pooling layer to the network to effectively reduce the dimensionality of the data, as shown in **Fig. 5**, the obvious loss decline curve becomes quite smooth, and the training effect is good. The accuracy of the test on the training set is 95%.

From this point of view, when performing face data identification, using convolutional layers and pooling layers to reduce the dimensionality of the picture can effectively prevent over-fitting during the learning process.

Experiments show that when the positive and negative samples are very different, the network cannot learn very well, resulting in an accuracy of only 66.7%. Therefore, the method of sample equalization and three local features are used to correct the results of the overall facial features, and finally the training set. The correct rate on the above is 99.4%, and the accuracy on the test set is 88%.

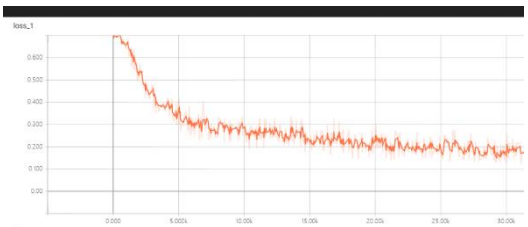


Fig. 5. Using convolutional layer and pooling layer

4. Conclusions

The main goal of this paper is to enhance the robustness of face recognition in all aspects. In order to achieve this goal, the design of face recognition is adopted to recognize whole and partial faces. On this basis, in order to eliminate the Haar-level classification in the feature extraction has an error of training a model for occlusion to solve the condition where the face cannot be recognized in some cases. We construct a Siamese network system, using the data set to construct positive and negative samples, adjust the ratio of positive and negative sample pairs, adjust training parameters, and use multi-channel local features to modify the overall facial features for the face identification algorithm strategy to obtain a significantly improved recognition accuracy.

References

[1] Shang-Hung Lin,Sun-Yuan Kung.Face recognition/detection by probabilistic decision-based neural network[J].IEEE Transactions on Neural Networks,1997,8(1).

[2] Taigman Y, Yang M , Ranzato M , et al. DeepFace: Closing the Gap to Human-Level Performance in Face Verification[C]// IEEE Conference on Computer Vision & Pattern Recognition. IEEE Computer Society, 2014.

[3] Chopra, Sumit & Hadsell, Raia & Lecun, Yann. (2005). Learning a similarity metric discriminatively, with application to face verification. Proc. Computer Vision and Pattern Recognition. 1. 539- 546 vol. 1. 10.1109/CVPR.2005.202.

[4] Parkhi O M , Vedaldi A , Zisserman A . Deep Face Recognition[C]// British Machine Vision Conference 2015.

[5] Y . LeCun and F. Jie Huang. Loss functions for discriminative training of energy-based models. AI-stats, 2005.

Efficient Object Counting Method for Object Detection and Tracking Services

Youngchan Lee[1], SoYeon Lee[1], Seokhoon Kim[2] and Dae-Young Kim[1*]
[1]School of Computer Software, Daegu Catholic University
Gyeongsan, South Korea
[2]Department of Computer Software Engineering, Soonchunhyang University
Asan, South Korea
[e-mail: {dldudcks1779, lsy5918647}@naver.com, seokhoon@sch.ac.kr, kimdy81@cu.ac.kr]
*Corresponding author: Dae-Young Kim

Abstract

Recently, object detection and tracking have become popular image-based services using deep neural network algorithm. In particular, the object detection and tracking are applied to movement pattern analysis of people or vehicles. This movement pattern analysis can be used for expansion of various applications for smart city. However, because of an inaccurate counting method by bounding box variation of objects, the efficiency of movement pattern analysis can be decreased. Therefore, in this paper, we propose improved counting method for object counting. Through the proposed method, the object counting is more accurate, and this can lead to improved efficiency for the movement pattern analysis.

***Keywords*:** Volumetric medical image, image segmentation, surface reconstruction, Augmented Reality

1. Introduction

Object detection and tracking using deep learning has been applied to various applications. In particular, movement data of people or vehicles, which is the result of the object detection and tracking, is used for decision making in various areas. For example, it can be usefully used in the problems such as analyzing the operation status of public transportation, identifying rescue targets, and evacuation routes in disaster, etc [1].

In general, the object detection and tracking is used to find objects and collect information. The information is analyzed and then applied to application services [2]. Deep learning technolgy acquires object recognition through learning input images. It obtains a bouding box for object as a learning result. Although the object recognition occurs for the same object, the bounding box size can be changed because of environmental changes such as color, brightness, and frequency in image frames. When objects are moving, the bounding box size changes very irregularly [3]. Due to this reason, when object counting by tracking is performed, object position is also irregularly changed and counting errors are generated.

Therefore, in this paper, we propose efficient object couting method to reduce the counting errors. The proposed method is applied to movement recognition of people and vehicles.

This research was supported by the KIAT (Korea Institue for Advancement of Technology) grant funded by the Korea Government(MOTIE: Ministry of Trade Industry and Energy) (No. P0014176).

2. Related work

2.1 Object detection and tracking

Object detection is composed of two procedures: classification and localization. Classification is recognizing what an object is. Localization is finding a position of an object using a bounding box [4].

Object tracking is calculated using the path approximation of the detected object in image frames. The position of the detected object is obtained through the center coordinates of the bounding box for the object. That is, moving path of the object is compared to the previous frame and recognize the same object. If the object is the same, tracking is performed [5].

2.2 MobileNet SSD

MobileNet is a kind of convolutional neural network (CNN) algorithms. It has less response time than other CNN algorithms because of reducing parameters in the neural network [8]. In addition, a single shot multibox detector (SSD) allows to perform learning for only one image to recognize an object. Through multi scale feature maps, high accuracy results are obtained without position estimation and resampling of the input images. Thus, it solves the trade-off problem of learning speed and accuracy. It uses non-maximal suppression (NMS) algorithm that remains high reliability objects and removes others [9]. Therefore, for fast and efficient object detection, MobileNet SSD, which combines MobileNet and SSD, is used.

3. System Implementation

3.1 Centroid tracking algorithm

The center coordinate of an object is caclulated from coordinates of the bounding box of detected object in a single frame. Unique ID is assigned to the center coordinate. As the object moves, the center coordinate in new frame changes and new objects may appear. The object with the shortest euclidean distance between new and previous coordinate is considered as the same object in the previous frame. Thus, it has the same ID.

$$d_E(p,q) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (1)$$

In an image frame, the euclidean distance means a distance between two points: (x_1, y_1) and (x_2, y_2) as represented in Eq.(1) [10].

The object movement is judged by the center coordinates with the minimum euclidean distance. If there are new center coordinates without the minimum euclidean distance, they becomes new objects and new IDs are assigned to their center coordinates. This procedure is repeated for each new image frame to track the detected object. That is, this algorithm is a multi-step process that tracks an object by calculating the Euclidean distance between the center coordinates of the objects in the previous frame and the next frame. **Fig. 1** shows the object tracking example.

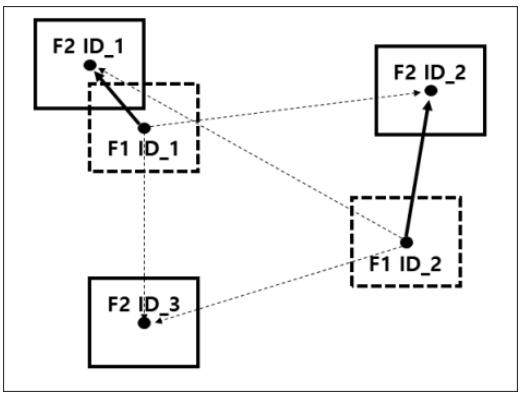


Fig. 1. Object tracking example

3.2 Counting algorithm

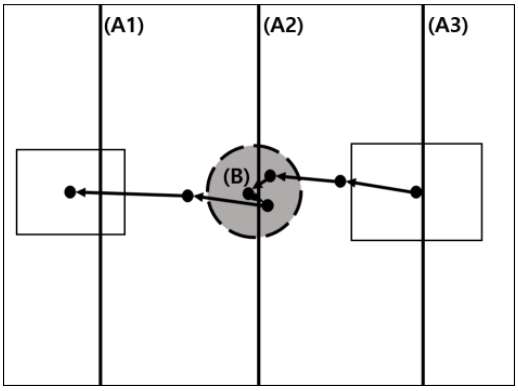


Fig. 2. Object tracking and counting

For object counting, bsaed on the central vertical line (A2 in **Fig. 2**), it is determined whether the starting point (center coordinate) of the object is located on the left or on the right. If the starting

point of an object is on the right side of A2 line in **Fig. 2** and the center coordinate of the moved object is on the left side of A1 line in **Fig. 2**, it is determined that it has moved to the left and counted. As the same manner, If the starting pont of an object is on the left side of A2 line and the center coordinate of the moved object is on the right side of A3, it is determined that it has moved to the right and counted.

In the object detection in video, the colors, brightness, and object size change every frame so that the size of the bounding box becomes irregular. This leads to inaccuracy center coordinate of the detected object. For example, in **Fig. 2**, the object B continues to move to the left, but it can be shown as moving to the right in some position. Thus, when a moving object is counted using only one vertical line, a counting error such as duplicate counting may occur. However, the proposed method avoids counting errors by using three vertical lines.

4. Evaluation

Table 1 represents the abscissa of the object B in the **Fig. 2** when it moves to the left. As mentioned earlier, it is shown that the object moves to the right at frame 4. In this case, the existing counting method performs an incorrect object count. In the experiment of the **Fig. 3**, the existing system counts three even though two people move to the left. In the vertical line A2, the counting error occurs.

Table 1. Abscissa of the object in **Fig. 2** when it moves to the left.

frame	1	2	3	4	5	6	7
abscissa	183	180	173	176	173	170	167

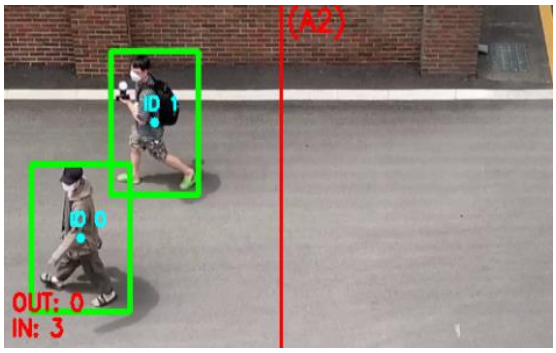


Fig. 3. The existing counting method

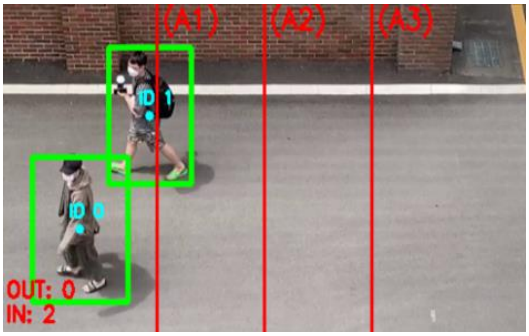


Fig. 4. The proposed counting method

In **Fig. 4**, the proposed counting method is applied. With employing three vertical lines, the proposed method improves counting accuracy. Thus, as shown in **Fig. 4**, when two people pass by, the proposed method counts exactly as two. The proposed method can eliminate counting errors by couting when the center coordinate pass through two vertical lines.

5. Conclusions

The bounding box is created as a result of object detection. It is variable in size due to environmental changes such as color and brightness in each frame. Thus, the center coordinate of the object becomes irregular. This causes an error in the counting mechanism of object tracking. In the proposed method, three vertical lines are used for counting. It imporves the counting accuracy even with irregular changes in the center coordinate.

References

[1] H. Lee, J. Lee, M. Kim, H. Nam, H. Mun and Y. Lee, “Measurement the floating population of the university using raspberry pi,” in *Proc. of Korea Computer Congress 2017*, Vol. 44, No. 2, pp. 2092-2094, 2017

[2] K. Kim, J. Heo and S. Kang, “Towards real-time multi-object tracking in CPU environment,” *Journal of Broadcast Engineering*, Vol. 25, No. 2, pp. 192-199, 2020.

[3] P. Lee and Y. Kim, “Advanced bounding box prediction with multiple probability map,” *Journal of Korea Society of Computer and Information*, Vol. 22, No. 12, pp. 63-68, 2017.

[4] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu and M. Pietikäinen, "Deep learning for generic object detection: A survey," in *arXiv:1809.02165 (Available: http://arxiv.org/abs/1809.02165)*, 2018.

[5] I. Pathan and C. Chauhan, "A survey on moving object detection and tracking methods," *International Journal of Computer Science and Information Technologies*, Vol. 6, No. 6, pp. 5212-5215, 2015.

[6] G. Jang and J. Park, "Visual object recognition based on deep neural networks implemented by CAFFE," in *Proc. of Communications of the Korean Institute of Information Scientists and Engineers*, Vol. 33, No. 8, pp.49-64, 2015.

[7] J. Lee, H. Kim, K. Bae and Y. Eo, "Detection of occlusion region using COCO data set," in *Proc. of Korean Society for Geospatial Information Science*, pp. 427-428, 2018.

[8] A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," in *arXrv: arXiv:1704.04861 (Available: https://arxiv.org/abs/1704.04861)*, 2017.

[9] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.Y. Fu and A.C. Berg, "SSD: Single shot multibox detector," in *Proc. of European Conference on Computer Vision (ECCV 2016)*, pp. 21-37, 2016.

[10] Y. Liu, H. Baek, S. Hong, J. Park and S. Kim, "User authentication service model using euclidean distance based on encryption," *Journal of Knowledge Information Technology and Systems*, Vol. 13, No. 6, pp. 695-703, 2018.

Dialogue Agent for Parent Management Training

Liguo Zhang[1], Joon-Young Paik[1,2,*]
[1] School of Computer Science and Technology, Tiangong University
399 Xiqing District, Tianjin, China
[2] Tianjin International Joint Research and Development Center of Autonomous Intelligence Technology and Systems, Tianjin, China
[*corresponding author: pjy2018@tiangong.edu.cn]

Abstract

The way of ADHD behavior treatment mostly is off-line, which need to communicate with therapist face-to-face. However, the parents aspire to get effective introduction in time while deal with children problem. Automatic approach of natural language-based interaction can satisfy this naturally demand well. In this paper, we make a first step towards a dialogue agent for parent management training whose children with ADHD. Specifically, we first build a dataset which is collected from real world situation where parents of children with ADHD consult therapists about how to deal with behavioral problems of their children. Then we propose dialogue system offering parent management training, which can converse with parents to collect additional behavior information beyond their self-report. Evaluations conducted in real-world settings show that additional behavior information extracted from the conversation can greatly improve the success ratio of the dialogue and make a better childhood ADHD therapy.

***Keywords*:** Dialogue Agent, ADHD, Parents Management Training

1. Introduction

Parent Management Training (PMT) is considered a best-practice for treating childhood ADHD [1]. On the one hand, parents of children with ADHD have a desire to obtain efficient introduction to deal with children behavior problems no matter whenever or wherever their children make an unreasonable activity. Under this demand, we propose a novel dialogue agent (DA) to support parents to make a correct and appropriate management. Especially, dialogue agent will play an important role for some parents who have their children got behavioral treatment and still be willing to have a consultant from therapist in order to master the knowledge and precautions for get along well with their children. On the other hand, Task-oriented dialogue system has been successfully applied to various application domains in recent years, such as ticket booking [2], online shopping [3], restaurant searching[4], and automatic diagnosis [5, 6]. We believe that applying DA in parent management training domain will help increase the convenience of families that have children with ADHD. Although having amount of data will facilitate the behavior treatment, the collection of information about children with ADHD by human labor is time-consuming. It is necessary to reduce the information collection cost by adopting an efficient automatic approach.

However, we face two challenges in applying DA in parent management training. The first one is that the lack of annotated behavior dialogue dataset. Another is no available DA framework for parent management training. By addressing these two problems, we make the first move to build a dialogue agent facilitating automatic information collection and parent management training whose children with ADHD. The

This work was supported by the Tianjin Science and Technology Program under Grant 19PTZWHZ00020.

contribution of this work is two-fold:

- We annotate the first ADHD behavioral management dataset for dialogue agent that consists of two parts, one is self-reports from parents and the other is conversational data between parents and therapist.

- We propose a dialogue agent framework to interact with people automatically. Experiment results on real situation show that our dialogue agent is able to improve the dialogue success ratio and make a better behavioral management automatically.

2. Knowledge Representation

Dataset is collected from group chatting in a Chinese group chatting application *WeChat*. Usually, parents would provide a self-report to therapist for describing the basic information of their ADHD children. Then a therapist will communicate with parents to get another information and give an explicit management solution based on both the self-report and the conversational data. An example is shown in **Table 1**. As we can see, the therapist can obtain additional information in conversion beyond self-report. In addition, we can also obtain the feedback afterwards to evaluate the efficiency of recommend solutions.

Table 1. An example of a user record

Self-report
I brought my child to hospital yesterday. He refused to have the blood tested. It has been four months for medication. The doctor insists on the blood test. What should I do?
Conversation
Therapist: How old is he/she? Parent: Seven years old. Therapist: Try the responsive listening and reward policy Parent: I have told him that I will buy him some delicious food after the blood test.
……

Three annotators is inveted to label the behavior phrases in both self-reports and conversational data, all of them are therapists and have rich experience in behavior treatment. The annotation is performed in two steps, namely behavior extraction and behavior classification.

2.1 Behavior Extraction

BIO (begin-in-out) schema is used for behavior identification. Each word is assigned a label of ``B", ``I" or ``O". In addition, we add a label of Intent to indicate the sentence-level intent. Also, each extracted behavior expression is factored into inform slot values indicating the whole process of the event. Two guidelines for the self-report and the conversational data are created to improve the annotation agreement between annotators. 1) Each record is annotated by at least two annotators. 2) Any inconsistency would be further judged by the third one.

2.2 Behavior Classification

Medical experts will mark each behavior expression with the classification after behavior expression identification. **Table 2** shows some phrases that describe behavior in the example and some related classification. The statistics of dataset is presented in **Table 3**.

After behavior extraction and classification, there are 6 unique problems category identified. In order to reduce the size of action space of the DA, only 18 behavior classifications with a frequency greater than or equal to 10 are kept. Samples are then generated, called user goal. **Table 4** shows an example of user goals.

Table 2. An example of behavior classification

Problem Category	Behavior Classification	Behavior Description
Behavior training	Procrastination	late for school

Table 3. The summary of dataset

Problem Category	Behavior Classification
Developmental	Study Hard
	attention
Behavior Training	procrastination
	Social
	Habit
	Interesting
Psychology	Stress
	Fear
Information Consulting	common consultation
	no-common consultation
Medical Treatment	Pharmacodynamic Effect
	Medical Treatment
	Medication Time
	Side Effect
Parents Confusion	Camp

Life
Book Selection
Diagnosis Result

Table 4. An example of user goal

```
{
  'solution': Reword policy, responsive
              listening
'request_slot':{
          'solution':UNK   }
'inform_slot':{
          'location': Hospital
        'behavior':No Blood Test }
}
```

3. Proposed Framework

The DA we proposed is consist of three components, namely Natural Language Understanding (NLU), Dialogue Management (DM) and Natural Language Generation (NLG). NLU is used for recognizing the user intent and slots with values from utterances; DM tracking the dialogue states and taking system actions; NLG generating natural language given the system actions. User simulator is applied for supporting train and interact with the dialogue system automatically and naturally. **Fig. 1** shows the framework.

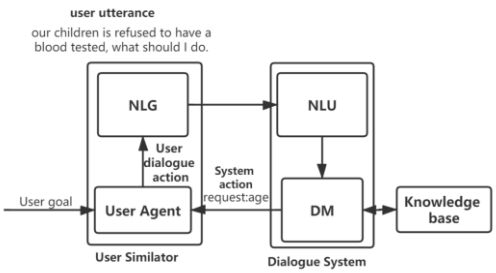


Fig. 1. Dialogue System Framework

3.1 User Simulator

The user simulator takes a user goal from the dataset randomly. User goal includes three parts: *inform_slots*, *request_slot*, and *solution_slot*. In details, *Inform_slot* provides the user's describing information, *request_slot* indicates something that the user wants to consult, and *solution_slot* provides some efficient treatments. At each turn t, the current user state $s_{u,t}$ and the previous agent action a_{t-1} is given to the user simulator; then it generates a user action and transits into the next user state $s_{u,t+1}$. In the whole process, the user simulator maintains a compact and stack-like representation called *user agenda*, where the user state s_u is factored into an agenda A [7] and a goal G, noted as $s_u = (A, G)$. During the dialogue, the goal G ensures that the user behaves in a consistent, goal-oriented manner. And the agenda contains a list of children behavior to track the progress of the conversation.

The user simulator will initialize a conversation by post a user action $a_{u,1}$, which contains requested *solution_slot* and all *inform_slot* values; then agent will check out knowledge base according to given *inform_slots* information. If all constrains is satisfied to a specific type of behavior classification, agent will give corresponding management solution. The dialogue session will be terminated as successful by the user. Otherwise, the dialogue session will be recognized as failed if the agent gives unsuitable management solutions or the dialogue turn reaches the maximum dialogue turn T. **Table 5** shows one success and one failure dialogue examples

Table 5. Two sample dialogues

User goal:	**User goal**:
{ 'solution_slot: Reward policy, Responsive listening 'request_slot': {solution: UNK} 'inform_slot': {'location':hospital; 'behavior': no blood analysis;} }	{ 'solution_slot': Mind map 'request_slot':{solution: UNK} 'inform_slot':{Grade: two grade; Behavior: have difficult in reciting; }
agent	**agent**
User: My son refuses to blood analysis.	User: Our children has difficult in reciting.
Agent: Have you discussed this with him at home?	Agent: Reciting need a good memory.
User: Yes, we are in hospital now.	User: He has been a student of two grade.
Agent: Responsive listening.	Agent: Children will be grow up.
User: Wow. That helps a lot! Thanks.	User: Do you know how to deal with that.
Agent: You are welcome.	Agent: You need to take care of your children.
Success	***Failure***: Dialogue turn more than maximum

3.2 Nature Language Understanding

The NLU is mainly implemented with a bidirectional LSTM[8], which captures temporal relationships within the sequence. As mentioned above, we have labeled the user utterance in stage of behavior classification, we use this dataset to train the model.

$$X = w_1, w_2, \dots, w_n \quad (1)$$
$$Y = s_1, s_2, \dots, s_n, i_m \quad (2)$$

Where X is the input word sequence and Y contains the associated slots s_k, and the sentence-level intent i_m.

$$H = biLSTM(X) \in \mathrm{R}^{n*d_{emp}} \quad (3)$$

Take the process of encoding a sequence into account with respect to a particular slot s. Let n denote the number of words in the sequence, d_{emp} the dimension of the embeddings, and $X \in \mathrm{R}^{n*d_{emp}}$ the word embeddings corresponding to words in the sequence.

3.3 Dialogue Management

As the same of traditional DM frame, the DM includes two stages, dialogue state tracking and policy learning, as shown in **Fig. 2**.

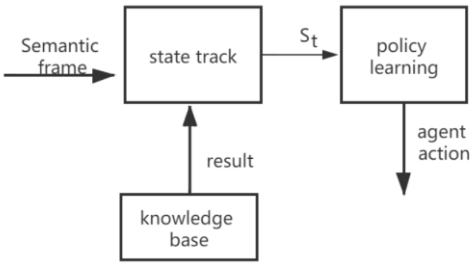


Fig. 2. Dialogue Management Module

3.3.1 Dialogue State Tracking

When the LU symbolic output, such as request (solution; behavior = late for school;), is passed to state tracker, three steps are performed one by one. Firstly, a symbolic query is formed to interact with the knowledge base to retrieve the available results; then based on the available results from the knowledge base and the latest user dialogue action, the state tracker will update the state s_t; finally, the state tracker will prepare the state representation s_t for policy learning.

3.3.2 Policy Learning

The state representation for the policy learning consist of the latest user action (e.g. request (solution; behavior=late for school), the latest agent action (request(age)), the available database results, turn information, and history dialogue turns. Considering the state representation S_t from the state tracker, the policy π is to generate the next available system action a_t according to $\pi(s_t)$.

3.4 Nature Language Generation

In nature language generation stage, the template-based NLG and the model-based NLG is combined to enhance generation performance, as shown in Figure 3. The model-based NLG is a sequence-to-sequence model. Given the dialogue acts, the model-based NLG will generate sentence sketch with slot placeholders via an LSTM decoder. Then a post-processing scan is performed to replace the slot placeholders with their actual values[2].

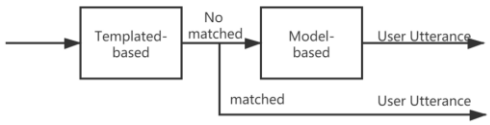


Fig. 3. Natural Language Generation

Firstly, the template-based NLG has a priority to match the user dialogue action according to predefined sentence templates. If matched, NLG will output the natural language texts generated by template-based NLG; otherwise, the utterance is generated by the model-based NLG. Applied the hybrid approach allow us to easily control the performance of NLG by providing template for sentences that the machine-learned model does not handle well.

4. Experiment and Results

4.1 Experimental Setup

To evaluate the performance of the proposed framework, we compare our model with baselines in terms of three evaluation metrics following [2] and [9, 10], namely, success rate,

average reward and the average number of turns per dialogue session.The baselines include: (1) Random Agent: At each turn, the random agent takes an action randomly from the action space as the response to the user's action. (2) Rule-based Agent: The rule-based agent takes an action based on handcrafted rules.

80% of the dataset is used for training and 20% for testing. The max dialogue turn T is 22. If a dialogue is successed, A positive reward of +44 is given to the agent, otherwise -22 reward is given to a failure one. In addition, step penalty of -1 reward is applied for for each turn to encourage shorter dialogues. Conditioned on the current dialogue state s_t, the agent will inform solutions if all the known slots value related are detected. If no solution can be given, the agent will select one of the left solutions randomly to offer. The relations between solutions and behaviors are extracted from the annotated corpus in advance. In this work, only the solutions with high frequency are kept for each behavior problems so that the rule-based agent could inform a solution within the max dialogue turn T.

4.2 Results

Observing from the result can get a conclusion that extracted the additional information from conversation will improve the success ratio of dialogue for all the behavior problems. Table 6 shows the performance of these agents on testing set. Due to the large action space, the random agent performs badly. The rule-based agent outperforms the random agent in a large margin. This indicates that the rule-based agent is well designed.

Table 6. Performance of Three Dialogue System

Model	Success	Reward	Turn
Random Agent	0.07	-23.26	15.36
Rule Agent	0.33	-12.68	19.00
DL Agent	0.54	17.62	5.82

5. Related Work

The goal-based dialogue system has been under development in recent years. [4] creates a restaurant reservation dialogue system .As the further step, [2] build a movie booking system, in this system, excepting include general Nature Language Understanding (NLU), Nature Language Generation (NLG), Dialogue Management (DM), it also includes user simulator[11], which simulate user in the training process. In addition, [12] make a move to build a dialogue system for automatically diagnosis, which can converse with patient to collect additional symptoms beyond patient self-reports and make a better diagnosis. In order to address the key problem in DST that is the extraction of rare slot-value pairs that composes the state during each turn, [13] proposed the Global-Locally Self-Attentive Dialogue State Tracker (GLAD), which learns representations of the user utterance and previous system actions with global-local modules.

6. Conclusions

In this paper, we construct a dataset which derived from the dialogue text between real parents of ADHD and therapists. Experimental results on a real-world dataset show that the proposed model can improve the dialogue success ratio and offer a better parent management training. Our future work includes constructing a knowledge graph of ADHD treatment domain. Reinforcement learning is under consideration to be applied in DST, which used to address problem of extraction of rare slot-value pairs.

References

[1] ELKINS S R, BOND M A, CURTIS D F. Do Comorbid Oppositional Symptoms Predict ADHD Behavioral Treatment Outcomes? [J]. Child Psychiatry & Human Development, 2019, 50(6): 918-26.

[2] LI X, CHEN Y-N, LI L, et al. End-to-end task-completion neural dialogue systems [J]. arXiv preprint arXiv:170301008, 2017,

[3] YAN Z, DUAN N, CHEN P, et al. Building task-oriented dialogue systems for online shopping; proceedings of the Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, F, 2017 [C].

[4] WEN T-H, VANDYKE D, MRKSIC N, et al. A network-based end-to-end trainable task-oriented dialogue system [J]. arXiv preprint arXiv:160404562, 2016,

[5] TANG K-F, KAO H-C, CHOU C-N, et al.

Inquire and diagnose: Neural symptom checking ensemble using deep reinforcement learning; proceedings of the NIPS Workshop on Deep Reinforcement Learning, F, 2016 [C].

[6] TOU H, YAO L, WEI Z, et al. Automatic infection detection based on electronic medical records [J]. BMC Bioinformatics, 2018, 19(Suppl 5): 117.

[7] SCHATZMANN J, THOMSON B, WEILHAMMER K, et al. Agenda-Based User Simulation for Bootstrapping a POMDP Dialogue System; proceedings of the Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Companion Volume, Short Papers, Rochester, New York, F apr, 2007 [C]. Association for Computational Linguistics.

[8] CHENG J, DONG L, LAPATA M. Long short-term memory-networks for machine reading [J]. arXiv preprint arXiv:160106733, 2016,

[9] PENG B, LI X, GAO J, et al. Adversarial advantage actor-critic model for task-completion dialogue policy learning; proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), F, 2018 [C]. IEEE.

[10] PENG B, LI X, LI L, et al. Composite Task-Completion Dialogue Policy Learning via Hierarchical Deep Reinforcement Learning [J]. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, 2017,

[11] LI X, LIPTON Z C, DHINGRA B, et al. A User Simulator for Task-Completion Dialogues [M]. 2016.

[12] WEI Z, LIU Q, PENG B, et al. Task-oriented Dialogue System for Automatic Diagnosis; proceedings of the Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), Melbourne, Australia, F jul, 2018 [C]. Association for Computational Linguistics.

[13] ZHONG V, XIONG C, SOCHER R. Global-Locally Self-Attentive Encoder for Dialogue State Tracking [J]. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2018,

Performance Analysis the of Cerebrospinal Fluid (CSF) Cell Diseases

Soobia Saeed, Afnizanfaizal Abdullah
School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia
Skudai, Johor Bharu, 81310, Malaysia
[E-mail: soobiasaeed1@gmail.com, Afnizanfaizal@utm.my]
*Corresponding author: saeed.soobia@graduate.utm.my

Abstract

Medicinal photographs play a crucial role in tumour diagnosis as well as in the leakage of cerebrospinal fluid (CSF). In a similar way, MRI may be the regenerative imaging technology of segmentation control, which enables the angle sectional viewpoint of the body to examine the injured person, which gives experts convenience. The purpose of this research is to develop a structure that can classify the region of cancer damage or be quietly separated from tumours and non-tumors by using 4D image light field segmentation, which is accompanied by AI software modelling techniques, and to quantify the size of brain damage cells deep within CSF. Data is usually collected from the SVM Tool by using MATLAB included K-Nearest Neighbour (KNN) Algorithm. This research proposes a 4Dimentional modulation method which supervises the light field which can be used in the field of light editing. Depending on the user's feedback, the objective assessment of each ray is determined using the KNN to preserve the light fields for 4D frequency (redundancy).

***Keywords*:** Brain Tumor, MRI, Image segmentation, CSF, KNN

1. Introduction

The cerebrospinal fluid (CSF) is the fluid that passes through the ventricles (cavities or vacuums) of the brain and across the surface of the brain. One of the most complex neuro-surgical complications is CSF. CSF leakage is a condition that occurs when the CSF leaks in the dura or head by deformity and escapes through the nose or ear. CSF leakage is the after-effect of a dura gap or tear, which is the most remarkable meningitis sheet. The motives behind the hole or tear may impact the brain or breast of the head and the medical system. After the lower back segment, also called spinal cord or spinal anaesthesia, CSF slots can be created in the same way. Without a known cause a CSF leak may occur in a similar manner without restriction. The cells that make up these interfaces are also sites of extensive exchange mechanisms (transporters) that govern the entry and exit of the brain into a wide range of molecules. The secretion of cerebrospinal fluid through choroid plexuses that circulate through the ventricular system and interchange between the cerebrospinal fluid and the brain is an important mechanism for controlling the specific

composition of the interstitial fluid of the brain. To assess the effects of inflammatory conditions on the brain, both in adults and during development, understanding the complexity of barrier mechanisms is important. Cerebrospinal fluid exhaustion may occur through leakage, a shunt, inadequate generation, or unusually quick retention.

K- Nearest Neighbours, also known as K-NN, is part of the family of supervised learning algorithms, which means that a collection of tagged data is used by the researcher to predict the new data point group. The K-NN algorithm is a strong workbook frequently used as a reference point for more complex ones, such as the Artificial Neural Network (ANN) or the SVM. The algorithm K-NN can be quickly grasped and implemented. The K-NN algorithm is read through a full collection of data to find the neighbours nearest to identify the new data point. As a research model, clustering is used.

Clustering is a very powerful technique for the detection of similarities between various clusters or classes.

2. LITERATURE REVIEW

The researcher describes how the tumour is separated into the brain by partitioning the tumour region from the MRI images. There are many ways that brain tumours can be easily segmented. In any event, the identification of the brain tumour from MR images is a repetitive task. The segmentation technique is for the extraction of various tumour tissues. Notwithstanding cerebrospinal fluid (CSF), another case of elements, tumour, debasement, oedema flowing from the normal brain tissues, such as dark and white tissues cells. As demonstrated by the mull-over survey, most brain tumours are effectively identified using a district-based attractive magnetic resonance imaging technique from cerebrum MRI. The degree of precision needed for the classification of abnormalities, however, is not surprising [1-4]. The segmentation of the brain tumours requires several steps. Brain photos that segment physical brain consume a lot, so there are many difficulties in manual segmentation. The main aim of this paper is to present half and half of the community in this exploratory study. Consisting of the C-Mean Fuzzy Clustering and a set of level strategies (to take care of complex forms) and level set strategy (to take care of complex forms) is to classify the tumour 's right condition at an insignificant time of computation. Using this approach, the author thought that it takes 0.9412 seconds to recognize a tumour that is less contrary to the current algorithm, that is, the cross-cross[5-6], for a particular collection of images. The specialist said in this research paper that segmentation of the picture is a collection of basic research, as it plays an important role in the investigation and understanding. Segmenting the images and partitioning the images is one of the toughest and toughest jobs. Filtered records can also be broken down from time to time in advance. Additional supplementary archive processing methods, such as compression or processing pressure or rendering, etc., can be related in advance[7]. The result shows that if the parameters are correctly configured in the AI software state based on the results, the fuzzy Level Set Segmentation for a number of levels will segment the tumour; It depends on the fact that the hybrid technique is used in a local way when the seed point is chosen. By improving the correct tumour limit in the brain by arranging the tumour pixels using the Fuzzy C-Means clustering method, an accurate tumour at measurement time is not important. In health effect pictures for the programmed division of brain tumours, segmentation techniques may be used. Additionally, work can be completed to influence this semi-automatic to automatic segmentation, with the goal of automatically calculating the tumour dimensions[8-12] as well. More precisely, the investigator has shown that there is an excellent relationship between the approximate sizes of CSF sequences and the T1-based 3D methods. The analysis aimed to check the size of the CSF series. The researcher discusses potential studies which could also concentrate on processing additional CSF images to recognise regional brain atrophy. This will allow for the discovery of different forms of atrophy in the brain. This is a valuable addition considering that many disorders have different brain atrophy, e.g. frontal dementia[13-18]. Subsequent procedures can also be shortened by one minute in future trials to promote adoption in clinical practise. The CSF sequence can also be verified in clusters of patients with further brain defects and brain disorder changes. CSF data is given by the CSF series. CSF-T, for example, may be linked to partial oxygen [19-23]. Finally, there is a similar resolution for each MRI-CSF sequence and a very strong correlation with the automated 3D-based segmentation methods generated by the BPV and ICV segments. The short imaging time for the MRI sequence in the CSF, however, exceeds the feasible three-dimensional sequence in T1 (3D T1) where particular methods are used to perform segmentation. Both the MRI-CSF series has a similar resolution and a very strong correlation with T1: Automated system-based segmentation methods created for the BPV and ICV partitions. The short imaging time of the magnetic resonance imaging MRI series, however, exceeds the 3D-T series in which the segmentation is calculated using fixed methods [24-25].

3. 4D Light Field Tool (LFT) Segmentation

There are several ways to depict the phase of

segmentation of the four-dimensional field of light. In order to demonstrate rays in three-dimensional space, the investigator followed the Lumigraph method. The beam (ray) is defined by two intersection points with "u-v" and "x-y" in the tridimensional coordinates. The ray can be represented as a point at a distance of 4D, such as p = (u, v, x, y), and the intensity p is represented. The representation of a lumigraph can be transformed into multiple representations containing a view level, an image level, and vice versa. The U-V and X-Y models correlate with the planes of view and picture in a multi-point view, respectively. In this study , the researcher discusses the easy-to - understand multicomponent representation system.

4. Comparative Evaluation

After applying the assessment criteria, the evaluation criteria will be learned and explained, and the findings and results will then be presented in the current methods. The aim of this study is to rank and compare the design of the most prominent research methods in order to verify the ability of the prevalent test approaches to enhance the testing of the training data set. we trained several supervised machine-learning models.

4.1 Evaluation Results

The results of these comparisons are explained in this section and are based on semi-supervised machine learning as we work on 4 Dimensional Picture Segmentation Procedures. We're focused on CSF leakage because of brain cancer. The test findings are carried out by a dataset named "Malignant Brain Cancer with CSF Leakage" which consists of nine variables with 25 characteristics. In this analysis, the initial information set of 4-dimensional data is then broken randomly into practise and test sets will be 96.9 percent proportional.

5. Conclusions

The 4d approach proposed aims to improve the sensitivity of the segmentation process to KNN algorithm technique parameters and improve efficiency, as well as to increase the performance of long-distance values compared to other comparative methods. KNN greatly increases the efficiency of distant values, primarily due to two primary reasons, brain cancer and leakage of CSF. The first improvement is that a set of adjacent values with parameters of neighbourhood size (K=1-100) is built in the KNN algorithm. Building these values ensemble, KNN will provide the optimised results with the low-dimensional. The efficiency of the methods analysed is to focus on the four Evaluation Metrics (EM) calculated, namely sensitivity (sen), specificity (spe) ROC, and accuracy (Acc). The second improvement is that the KNN algorithm provides an accurate weighting for all far values, making the predictions much less sensitive compared to previous work to the neighbourhood size parameter.

References

[1] Altaf I, Vohra AH, Shams S. Management of cerebrospinal fluid leak following posterior cranial fossa surgery. Pak J Med Sci. 2016; 32:1439-1443.

[2] Green AL, Arnaud A, Batiller J, Eljamel S, Gauld J, Jones P, et al. A multicentre, prospective, randomized, controlled study to evaluate the use of a fibrin sealant as an adjunct to sutured dural repair [e-pub ahead of print]. Br J Neurosurg. 2014:1-7.

[3] Philipp Kickingereder, Fabian Isensee, Irada Tursunova, Jens Petersen, Ulf Neuberger, David Bonekamp, Gianluca Brugnara, Marianne Schell, Tobias Kessler, Martha Foltyn, Inga Harting, Felix Sahm, Marcel Prager, Martha Nowosielski, Antje Wick, Marco Nolden, Alexander Radbruch, Jürgen Debus, Heinz-Peter Schlemmer, Sabine Heiland, Michael Platten, Andreas von Deimling, Martin J van den Bent, Thierry Gorlia, Wolfgang Wick, Martin Bendszus, Klaus H Maier-Hein. (2019).Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: a multicentre, retrospective study. *The Lancet Oncology*, pp.1-12.

[4] Schievink, W. I., et al. (2016). "A classification system of spontaneous spinal CSF leaks." Neurology 87(7): 673-679.

[5] Sin AH, Caldito G, Smith D, Rashidi M, Willis B, Nanda A.(2019). Artificial intelligence helps to better assess treatment response of brain tumors. German Cancer

Research Center vol.(5):224-227.

[6] Badarnel, A.A., Nafadat, H. and Alraziqi A.M. 2012. A classifier to detect tumor disease in brain MRI brain images.ACM international conference on advances in social networks analysis and mining, IEEE.

[7] Algfoor et al. (2017). A new weighted pathfinding algorithms to reduce the search time on grid maps, Expert Systems and Applications.

[8] Junting Zhao, Zhaopeng Meng, Leyi Wei, Changming Sun, Quan Zou and Ran Su (2019) Supervised Brain Tumor Segmentation Based on Gradient and Context-Sensitive Features. Brain imaging method, pp 1–12.

[9] Wu W, Chen AY, Zhao L, Corso JJ (2014) Brain tumor detection and segmentation in a CRF (conditional random fields) framework with pixel-pairwise affinity and super pixel-level features. Int J Comput Assist RadiolSurg 9:241–253.

[10] Soobia Saeed, Afnizanfaizal Abdullah, NZ Jhanjhi," Investigation of a Brain Cancer with Interfacing of 3-Dimensional Image Processing", Indian Journal of Science & Technology, 2019.

[11] Su YeonLee , Min Kyung Ju , Hyun Min Jeon, EuiKyong Jeong, YigJi Lee,1 Cho Hee Kim, HyeGyeong Park, Song Iy Han , and Ho Sung Kang.(2018)." Regulation of Tumor Progression by Programmed Necrosis", Oxidative Medicine and Cellular Longevity Vol.28, pp.1-19.

[12] Soobia Saeed, Afnizanfaizal Abdullah, (2019)," Investigation of a Brain Cancer with Interfacing of 3-Dimensional Image Processing", International Conference on Information Science and Communication Technology (ICISCT),PP.1-12.

[13] Sajeel Chowdhary, MD, Sherri Damlo, and Marc C. Chamberlain, MD. (2017). "Cerebrospinal Fluid Dissemination and Neoplastic Meningitis in Primary Brain Tumors", journal of mofitt cancer centre, Vol. 24(1):1-16.

[14] Louis S. Prahl, Maria R. Stanslaski , Pablo Vargas , MatthieuPiel, and David J. Odde. (2018)." Glioma cell migration in confined microchannels via a motor-clutch mechanism", bioRxiv , pp.1-30

[15] Soobia Saeed, Afnizanfaizal Abdullah, NZ Jhanjhi," Implementation of Fourier Transformation with Brain Cancer and CSF Images", Indian Journal of Science & Technology, 2019.

[16] Yatsushiro. (2014) Correlation mapping for visualizing propagation of pulsatile CSF motion in intracranial space based on magnetic resonance phase contrast velocity images: Preliminary results, EMBC 2014

[17] J. Platt. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. Advances in large margin classifiers, vol.10 (3):61–74, 2014.

[18] A. Kowdle, S. N. Sinha, and R. Szeliski. Multiple view object co-segmentation using appearance and stereo cues. In Proc. ECCV, pp.64-65, 2012.

[19] Abdullah. (2013). An Improved Swarm Optimization for Parameter Estimation and Biological Model Selection, PLOS ONE

[20] Ahmet Kinaci , Ale Algra , Simon Heuts , Devon O'Donnell, Albert van der Zwan, Tristan van Doormaal, (2018), "Effectiveness of Dural Sealants in Prevention of Cerebrospinal Fluid Leakage After Craniotomy: A Systematic Review" ,World Neurosurg. , vol.118:368-376.

[21] Norman R. Saunders, Katarzyna M. Dziegielewska, Kjeld Møllgard , and Mark D. Habgood,(2018), " Physiology and molecular biology of barrier mechanisms in the fetal and neonatal brain", The Journal of Physiology, pp 5723–5756

[22] Sivan Gelb a, Ariel D. Stock b, Shira Anzi a , Chaim Putterman b , Ayal Ben-Zvi, (2018), "Mechanisms of neuropsychiatric lupus: The relative roles of the bloodcerebrospinal fluid barrier versus blood-brain barrier", Journal of Autoimmunity, Science Direct , USA, pp.1-11.

[23] Shawn Kant , Edward G. Stopa , Conrad E. Johanson , Andrew Baird and Gerald D. Silverberg, (2018), "Choroid plexus genes for CSF production and brain homeostasis

are altered in Alzheimer's disease", Fluids and Barriers of the CNS,vol.15(34):1-10.

[24] Wen-xuan Jiana,b , Zhao Zhangb , Shi-feng Chub , Ye Pengb , Nai-hong Chen, (2018), "Potential roles of brain barrier dysfunctions in the early stage of Alzheimer's disease", Brain Research Bulletin, vol.142, pp.360-367

[25] van der Kleij LA, de Bresser J, Hendrikse J, Siero JCW, Petersen ET, De Vis JB (2018) Fast CSF MRI for brain segmentation; Cross-validation by comparison with 3D T-based brain segmentation methods. PLoS ONE 13(4): e0196119.

Coronary artery calcium detection from heart CT scans using deep learning

Sungjin Lee[1], Beanbonyka Rim[1], JunSeob Kim[1], Xibin Jia[2] and Min Hong[3]
[1]Department of Software Conversions, Soon chun hyang University
Asan, South Korea
[2]Faculty of Information Technology, Beijing University of Technology
Beijing, China
[3]Department of Computer Software Engineering, Soonchunhyang University, Asan, South Korea
[e-mail: fijianwa@gmail.com, rim.beanbonyka@gmail.com, openfolder0220@gmail.com, jiaxibin@bjut.edu.cn, mhong@sch.ac.kr]
*Corresponding author: Min Hong

Abstract

The frequency of diseases caused by arteriosclerosis is also increasing due to the increase of the elderly due to the improvement of the living environment, and the death and morbidity rates from heart disease are also on a rapid increase. In order to determine the degree of coronary artery calcification, the most common method is to determine the calcification score through a heart CT scan. In order to prevent such calcification and to inform the patient through accurate judgment and calculation in advance, and to enable the patient to take care of his or her health, the calcification score for each coronary artery is judged while viewing chest pain CT manually. Because of this, the time to receive the results can vary widely from person to person. To prevent this, in this study, an experiment was conducted to see if it is possible to assist with reading after learning a total of 150 people using deep learning.

***Keywords*:** Coronary artery, Ct, deeplearning, heart, inception-resnet-v2

1. Introduction

The frequency of diseases caused by arteriosclerosis is also increasing due to the increase of the elderly due to the improvement of the living environment, and the death and morbidity rates from heart disease are also on a rapid increase [1]. Among these cardiac diseases, the cause of most ischemic heart disease is arteriosclerosis, and it occurs a lot in middle-aged to old age, but it has already begun in adolescence and is caused as a result of progression for decades [2]. Therefore, it is necessary to detect atherosclerosis before symptoms appear because there is a time when atherosclerosis occurs and ischemic heart disease occurs. Methods of discovering this include vascular ultrasonography, magnetic resonance imaging, measurement of blood lipid metabolism, and electron beam tomography. Here we decided to proceed with CT. To determine the degree of coronary artery calcification, the most common method is to determine the calcification score through chest pain CT scan. [3]. The heart arteries are arterial blood vessels that supply oxygen and nutrients to the heart muscle. The shape surrounding the heart is similar to a tube used in the head, so it is also called a coronary artery. This phenomenon in which wastes or calcium accumulate on the wall of the blood vessel of the heart artery and become hardened is

This research was supported under the framework of international cooperation program managed by the National Research Foundation of Korea(NRF-2019K1A3A1A20093097) and this research supported by National Key Research and Development Program of China(2019YFE0107800) and this research supported by Soonchunhyang University's research fund.…..

called calcification. This calcification is the cause of myocardial infarction, angina, and ischemic heart disease because the blood flow is not smooth as the cardiovascular wall is narrowed [4]. In order to prevent such calcification and to inform the patient through accurate judgment and calculation in advance, and to enable the patient to take care of his or her health, the calcification score for each coronary artery is judged while viewing chest pain CT manually. Because of this, the time to receive the results can vary widely from person to person. To prevent this, in this study, an experiment was conducted to see if it is possible to assist with reading after learning a total of 150 people using deep learning.

2. Inception Resnet V2

Convolutional networks have become popular in large scale image recognition tasks after Krizhevsky et al. [5]. Next important milestones were Network-innetwork [6] by Lin et al., GoogLeNet (Inception-v1) [7] by Szegedy et al. and VGGNet [8] by Simonyan et al. The Inception-Resnet-v2 was released from Google. It is a convolutional neural network (CNN) that achieves great accuracy on the ILSVRC image classification benchmark. It is based from Microsoft's ResNet papers [9,10]. Residual connection was introduced by He et al. in [11] in which have advantages of additive merging of signals for image recognition. We argue that residual connections are necessary for training very deep convolutional models.

3. Training data set and result

We collected the data from Soon cheon hyang university hospital in Cheonan. Approximately 2400 people's CT data was collected and the CT image was approx. 120,000 slices. All of the data has been de-identified. The specific calcium score was only written in result file. So, we detect the calcium score for each slice. Only approx. 3500 slices of data were used. We divide the group into calcium and non-calcium. Each group have 1,500 training data and validation data was 150. 10% of training data.

Table 1. Data set of paper

	Training	Validation
Calcium	1500	150
Non-Calcium	1500	150

On our trial we reshape the size into 299x299 and made three drop out layer to prevent overfitting, for training calcium and non-calcium, the spine nearly shows the similar HU value with calcium. Therefore, we use cardiac segmentation to use only heart point to avoid the bone part. **Fig. 3** shows the flow of segmenting the cardiac area.

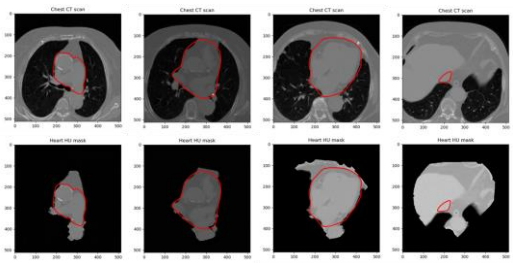


Fig3. Segmenting the cardiac area

The result graph is shown on **Fig. 4** and **Fig. 5**. The training accuracy was approx 58% but for the validation accuracy we got approx 62%. This result was the highest when we got result.

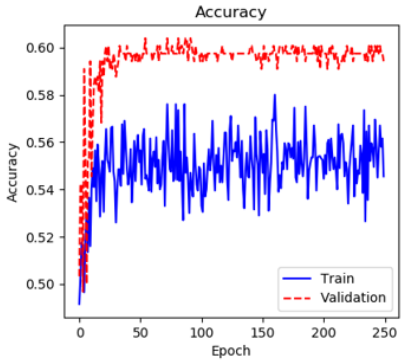


Fig 4. Last trial accuracy graph of result

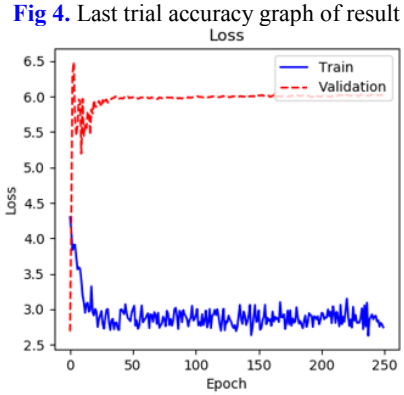


Fig 5. Last trial loss graph of result

4. Conclusions

In this paper, we designed a model to detect calcium in the heart. These models can save time to analyze CT images in the future, and can detect areas that people may have missed, thereby improving the quality of life for many people. However, it is still showing low accuracy in terms of accuracy and has not collected enough data. Because of this, there is a long way to directly or indirectly help humans, but based on these studies later, deep learning about the heart is conducted, and when such a heart is recognized, once again, learning to distinguish between a heart with an abnormality and a heart without an abnormality proceeds. Finally, we plan to develop a model that can provide detailed scoring for such a heart.

References

[1] Sakboonyarat, Boonsub, and Ram Rangsin. "Prevalence and associated factors of ischemic heart disease (IHD) among patients with diabetes mellitus: a nation-wide, cross-sectional survey." BMC cardiovascular disorders 18.1 (2018): 151.

[2] Stary HC. Evolution of atherosclerotic plaques in the coronary arteries of young adults. Arteriosclerosis 1983;3:471a.

[3] Becker CR, Knez A, Becker A, et al. Comparison of coronary calcium scoring by electron beam computed tomography and prospectively ECG-gated conventional computed tomography. Advances in Electron Beam Computed Tomography:1988:11.

[4] Arad Y, Spadaro L, Goodman K, et al. Predictive value of electron beam computed tomgraphy of the coronary arteries. 19-month follow-up of 1173 asymptomatic subjects. Circulation 1996;93:1951-53

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pages 1097–1105, 2012.

[6] M. Lin, Q. Chen, and S. Yan. Network in network. arXiv preprint arXiv:1312.4400, 2013.

[7] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1–9, 2015.

[8] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

[9] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[10] He, Kaiming, et al. "Identity mappings in deep residual networks." European conference on computer vision. Springer, Cham, 2016.

[11] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. arXiv preprint arXiv:1512.03385, 2015.

Deep Learning based Cardiac Arrest Prediction using Vital sign and Lab code

Minsu Chae[*], Sangwook Han[*], Geonil Yun[**], Hyo-Wook Gil[***], Min Hong[**], DoKyeong Lee[*], HwaMin Lee[**]

[*]Department of Software Convergence, Soonchunhyang University
[**] Department of Computer Software and Engineering, Soonchunhyang University
[***] Department of Internal Medicine, Soonchunhyang University Cheonan hospital
Chungcheongnam-do, South Korea
[e-mail: cmspr0@nate.com, sangwuk@sch.ac.kr, rjsdlf45@gmail.com, hwgil@schmc.ac.kr, mhong@sch.ac.kr, dooky606@daum.net, leehm@sch.ac.kr]
*Corresponding author: HwaMin Lee

Abstract

There are patients in hospitals who are hospitalized for a variety of reasons. We conducted a study on predicting cardiac arrest on patients at Soonchunhyang University Cheonan Hospital. We studied patients from 2016 to 2019. We used deep learning via the LSTM model and the GRU model. We check density of each input feature according to cardiac arrest. We compared only patients with vital signs and lab data. We removed DBP, SBP, BodyTemperature, AST, ALT, WBC, Creatinine, and Bilirubin variable density because there was little difference between those with cardiac arrest and other patients. We experimented with the LSTM model and GRU Model. In this paper, deep learning-based cardiac arrest prediction using vital signs and lab data has high precision. In particular, when using the GRU model, even if the cardiac arrest of 0 to 24 hours for each record is changed, it has a sensitivity of 60%.

***Keywords*:** Cardiac Arrest, GRU Model, LSTM Model, Deep Learning

1. Introduction

There are patients in hospitals who are hospitalized for a variety of reasons. Cardiac arrest is a disease in which the heart stops. If the response, such as CPR, is late, the body function is impaired, and in particular, brain damage can occur. The number of cardiac arrest patients in South Korea increased from 19480 in 2006 to 29262 in 2017[1]. The rapid response system in the hospital uses the Modified Early Warning Score (MEWS) to predict cardiac arrest patients[2]. However, since the sensitivity of MEWS is 20~60%, there are problems of low sensitivity and false alarms[3]. Accordingly, in this paper, we intend to predict cardiac arrest for patients in the hospital.

2. Related Works

2.1 Long Short-Term Memory (LSTM)

The LSTM model is a kind of Recurrent Neural Network (RNN) used in Deep learning. The RNN model has a long-term dependency

This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Korean government (MSIT) (No. NRF-2019M3E5D1A02069073) and supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2020-2015-0-00403) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation)

problem, the LSTM model solves the long-term dependency problem through a long-term memory unit and a short-term memory[4].

2.2 Gated Recurrent Unit (GRU)

The GRU model improves the problems of the LSTM model. LSTM model has a structure of Input Gate, Forget Gate, and Output Gate. The GRU model predicts similar results to LSTM while simplifying the equation with Reset Gate and Update Gate[5].

2.3 Prediction of cardiac arrest

Recently, studies for predicting cardiac arrest are in progress[3,6-7]. The most representative study is the Deep Learning-based Early Warning System (DEWS) [3]. For DEWS, only 8 hours of vital signs are used[3]. And, they use only blood pressure, SBP, respiration, and body temperature as vital signs[3].

3. Experiments and Result

3.1 Materials

We conducted a study on predicting cardiac arrest on patients at Soonchunhyang University Cheonan Hospital. We studied patients from 2016 to 2019. **Table 1** shows the number of patient data. We processed the sequence data into three-dimensional data.

Table. 1. Number of patients

Disease	Number of patients
Cardiac Arrest	1,154
Other	82,389

Table 2 shows the input features. MeasurementTime is a value that shows the measurement date and time using Timestamp. Age is the age when hospitalization. Gender is 1 for the men and 2 is for the women.

Table. 2. Input features

Feature	Describe
Gender	One is man, 2 is woman
Age	Age when hospitalzation
MeasureTime	Timestamp
DBP	Diastolic blood pressure (mmHg)
SBP	Systolic blood pressure(mmHg)
BodyTemperature	Body Temperature
Breath	Breath per minute
BloodPressure	Blood Pressure (mmHg)
Albumin	Albumin
AlbuminCheck	One has value, 0 is None
AST	Aspartate Aminotransferase
ASTCheck	One has value, 0 is None
ALT	Alanine Aminotransferase
ALTCheck	One has value, 0 is None
Bilirubin	Bilirubin
BilirubinCheck	One has value, 0 is None
Creatinine	Creatinine
CreatinineCheck	One has value, 0 is None
PLT	Count of platelet
PLTCheck	One has value, 0 is None
HB	Count of hemoglobin
HBCheck	One has value, 0 is None
WBC	Count of white blood cell
WBCCheck	One has value, 0 is None
CardiacArrest	One is patients with cardiac arrest, 0 is partients of other

We defined the data range, as shown in Table 3, to ignore abnormal values in the vital signs.

Table. 3. Range of vital signs

Vital sign	Range
DBP	>= 30 and <=300
SBP	>= 30 and <=300
BodyTemperature	>= 30 and <=45
Breath	>= 3 and <=60
BloodPressure	>= 30 and <=300

Doctor measure vital signs and lab codes according to the patient's condition. We replaced the missing values in the vital sign and lab code with the last measured values.

3.2 Result of LSTM model

Fig. 1 is the architecture of a cardiac arrest prediction model using three layers of the LSTM model.

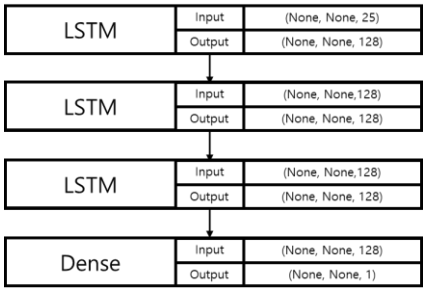


Fig. 1. Architecture of cardiac arrest model using three layers of the LSTM model

Fig. 2 is a graph of whether the training data and the predicted data check cardiac arrest. The blue line is the actual value, and the orange line is the predicted value. One means cardiac arrest, zero means other cases.

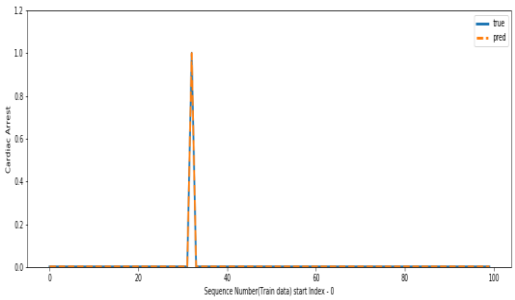


Fig. 2. A graph of training data and the predicted data via the LSTM model

Fig. 3 is a graph of whether the training data and the predicted data check cardiac arrest. The blue line is the actual value, and the orange line is the predicted value. One means cardiac arrest, 0 means other cases.

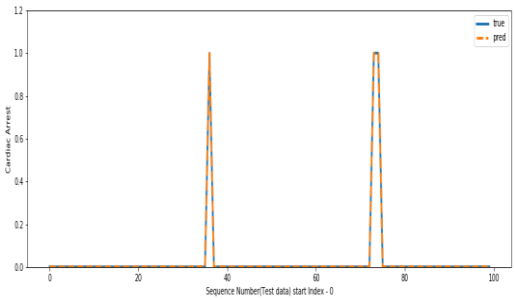


Fig. 3. A graph of test data and the predicted data via the LSTM model

Table 4 shows the results of the confusion matrix of test data using the LSTM model.

Table. 4. Confusion matrix of test data using the LSTM model

		Prediction	
		Cardiac arrest	No cardiac arrest
Actual	Cardiac arrest	5,184	0
	No cardiac arrest	0	355,752

Table 5 shows the result of the performance evaluation of the LSTM model. 5

Table. 5. Performance evaluation of the LSTM model

Accuracy	1.0
Precision	1.0
Recall	1.0

3.3 Result of GRU model

Fig. 4 is the architecture of a cardiac arrest prediction model using three layers of the GRU model.

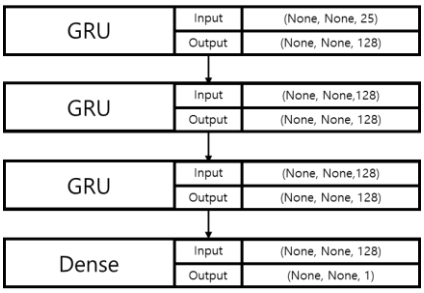


Fig. 4. Architecture of cardiac arrest model using three layers of the GRU model

Fig. 5 is a graph of whether the training data and the predicted data check cardiac arrest. The blue line is the actual value, and the orange line is the predicted value. One means cardiac arrest, 0 means other cases.

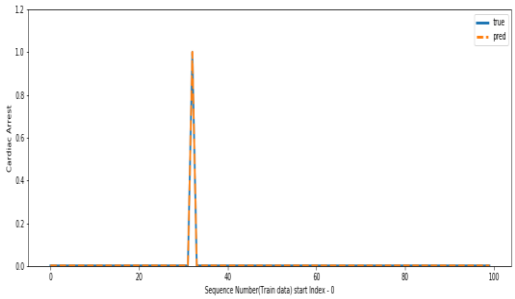


Fig. 5. A graph of test data and the predicted data via the GRU model

Fig. 6 is a graph of whether the training data and the predicted data check cardiac arrest. The blue line is the actual value, and the orange line is the predicted value. One means cardiac arrest, 0 means other cases.

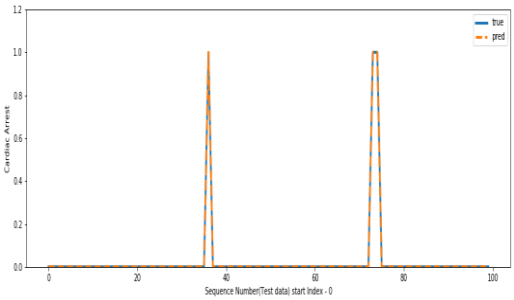


Fig. 6. A graph of test data and the predicted data via the LSTM model

Table 6 shows the results of the confusion matrix of test data using the GRU model.

Table. 6. Confusion matrix of test data using the GRU model

		Prediction	
		Cardiac arrest	No cardiac arrest
Actual	Cardiac arrest	5,184	0
	No cardiac arrest	0	355,752

Table 7 shows the result of the performance evaluation of the GRU model.

Table. 7. Performance evaluation of the GRU model

Accuracy	1.0
Precision	1.0
Recall	1.0

3.4 Density of input feature and experiment results using only some input features

Fig. 7 shows the density of each input feature according to cardiac arrest. We compared only patients with vital signs and lab data. We removed DBP, SBP, BodyTemperature, AST, ALT, WBC, Creatinine, and Bilirubin variable density because there was little difference between those with cardiac arrest and other patients. We experimented with the model structures used in sections 3.2 and 3.3.

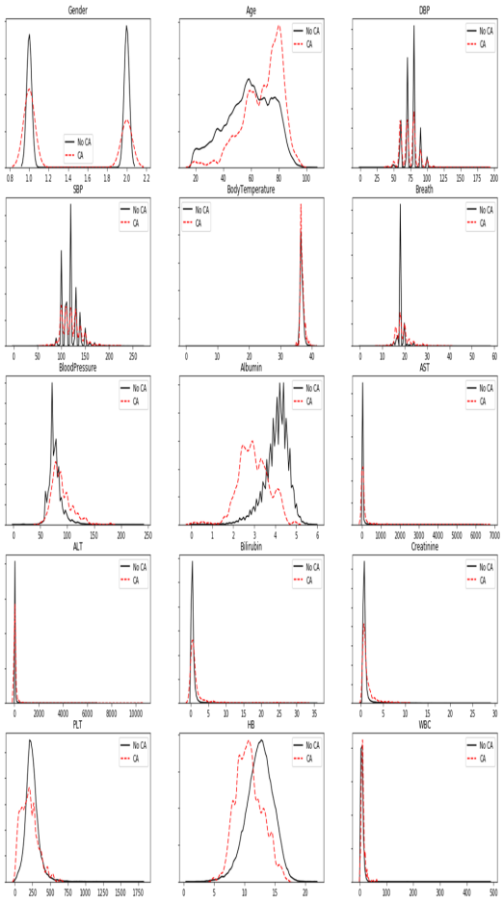


Fig. 7. The density of input features

Table 8 shows the result of a confusion matrix of test data using only some input features using the LSTM model.

Table. 8. Confusion matrix of test data using the LSTM model with only some input features

		Prediction	
		Cardiac arrest	No cardiac arrest
Actual	Cardiac arrest	3,371	1,813
	No cardiac arrest	1,740	353,940

Table 9 shows the result of a confusion matrix of test data using only some input features using the GRU model.

Table. 9. Confusion matrix of test data using the GRU model with only some input features

		Prediction	
		Cardiac arrest	No cardiac arrest
Actual	Cardiac arrest	4.452	1,732
	No cardiac arrest	1,366	355,752

Table 10 shows the performance evaluation of using only some features using the LSTM model and the GRU model.

Table. 10. Performance evaluation of the LSTM model and the GRU model with only some input features

Model	Accuracy	Precision	Recall
LSTM	0.990154	0.659558	0.65027
GRU	0.991449	0.71648	0.665895

3.5 Experimental results of dependence on cardiac arrest feature

We performed deep learning without the feature of cardiac arrest but did not get any meaningful results. We value of cardiac arrest was changed to test whether the dependence of cardiac arrest or not. We conducted experiments by randomly extracting 50 cardiac arrest patients and 50 other patients without overlapping. We changed the cardiac arrest from the measurement time to the specified time from 1 to 0 and 0 to 1.

Table. 11. Result of performance evaluation dependence on cardiac arrest feature

Model	Time	Accuracy	Precision	Recall
LSTM model	0~8	0	0	0
	0~16	0	0	0
	0~24	0	0	0
	0~32	0	0	0
	0~40	0	0	0
	0~48	0	0	0
	0~56	0	0	0
	0~64	0	0	0
	0~72	0	0	0
LSTM model with onpy some features	0~8	0.007222	0	0
	0~16	0.007919	0	0
	0~24	0.007361	0	0
	0~32	0.007222	0	0
	0~40	0.007361	0	0
	0~48	0.007639	0	0
	0~56	0.007222	0	0
	0~64	0.007083	0	0
	0~72	0.007361	0	0
GRU model	0~8	0.824306	0.832716	0.811667
	0~16	0.705	0.710256	0.6925
	0~24	0.598889	0.601598	0.585556
	0~32	0.485417	0.485047	0.473056
	0~40	0.373472	0.369522	0.358333
	0~48	0.262639	0.256344	0.249722
	0~56	0.154583	0.146029	0.1425
	0~64	0.041389	0.028016	0.027222
	0~72	0	0	0
GRU model with onpy some features	0~8	0.259583	0.083293	0.048056
	0~16	0.223194	0.006439	0.003611
	0~24	0.14625	0	0
	0~32	0.089861	0.006682	0.005556
	0~40	0.058333	0.006211	0.005556
	0~48	0.0275	0.001465	0.001389
	0~56	0.008333	0	0
	0~64	0.000139	0	0
	0~72	0	0	0

We confirmed that there is a dependency on the cardiac arrest feature through the results of Table 11.

4. Conclusions

There was a problem with low sensitivity in predicting cardiac arrest. In this paper, deep learning-based cardiac arrest prediction using vital signs and lab data has high precision. In particular, when using the GRU model, even if the Cardiac Arrest of 0 to 24 hours for each record is changed, it has a sensitivity of 60%. In the future, to solve the problem depends on the cardiac arrest feature, we intend to develop a model that calculates the cardiac arrest score using vital signs and lab code and predicts cardiac arrest through the cardiac arrest score. Also, it is intended to remove unnecessary input features by studying vital signs and lap data related to cardiac arrest.

References

[1] Korea Disease Control and Prevention Agency, “2006-2017 Acute Cardiac Arrest Survey Statistics”, 2018.

[2] Subbe, C. P., et al. “Validation of a modified Early Warning Score in medical

admissions." *QJM: An International Journal of Medicine*, Vol 94. No.10, pp. 521-526, 2001.

[3] Kwon, Joon- myoung, et al. "An algorithm based on deep learning for predicting in- hospital cardiac arrest." *Journal of the American Heart Association*, Vol 7. No.13, e008678, 2018.

[4] Gers, Felix A., Jürgen Schmidhuber, and Fred Cummins. "Learning to forget: Continual prediction with LSTM." in *Proc. of 9th International Conference on Artificial Neural Networks: ICANN '99*, pp. 850-855, 1999.

[5] Cho, Kyunghyun, et al. "Learning phrase representations using RNN encoder-decoder for statistical machine translation." in *Proc. of 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP2014)*, pp.1724-1734, 2014

[6] Park, Sung Bum, et al. "Developing a risk prediction model for survival to discharge in cardiac arrest patients who undergo extracorporeal membrane oxygenation." *International journal of cardiology*, Vol 177. No.3, pp. 1031-1035, 2014.

[7] Dumas, Florence, Wulfran Bougouin, and Alain Cariou. "Cardiac arrest: prediction models in the early phase of hospitalization." *Current opinion in critical care*, Vol.25. No.3, pp. 204-210, 2019

A Customer Demographics-aware Restaurant Recommendation System: A Deep Learning Approach

Afaq Muhammad, Gangseok Kim, Jie Xu, Yong-doo Jin, and Wang-Cheol Song
Department of Computer Engineering, Jeju National University
Jeju, South Korea
[E-mail : afaq24@gmail.com, kingiron@gmail.com]
*Corresponding author : Wang-Cheol Song

Abstract

Lack of information about the restaurant and the large number of restaurant choices turn into an obstacle to customers' requirements in selecting a restaurant. To this end, in this paper, we propose a recommendation system that can recommend a restaurant based on customer's demographics (age and gender) and rating given by other customers. We implement the proposed system that exploits deep learning techniques for age and gender recognition when given a customer's image as an input. After detecting age and gender of the customer, the system recommends a restaurant, to which the highest ratings were previously given by the customers of the same age group and gender to that of the currently visiting customer.

***Keywords*:** demographics-aware; recommendation system; restaurant; deep learning

1. Introduction

Humans are meant to derive satisfaction and pleasure from meals. So it is very important for them to have access of the best meals-serving places in their localities. Each individual has its own food preferences, which are mostly motivated by peers and same age-group members. For example young people are fond of going to fast food restaurants, coffee shops and mostly crowded places. In contrast, old people like calm and traditional food places. In addition, an individual's health also plays an important role in its food preferences. So, it is very complex to recommend a restaurant to a person without knowing the personality and health of the person. Few of the researchers focused on user's health parameters by installing sensors inside the restaurant to recommend specific food to a user [1,3]. However, they do not consider the user's demographics when recomminding a restaurant.

Recommendation systems work based on user's history. There are two ways of recommending, i.e., colaborative filtering for recomending based on similaraties between users, and content filtering based on the similaraties in contents. Nowadays Machine Learning (ML)/Deep Learning (DL) is widely used for recommendation, classification, prediction and recognition [6]. Usually, opting between the collaborative filtering or content filtering depends on the use-case. For example Youtube recommendation system works in both the modes, i.e., if a user uses Youtube without login credential and history, the recommender engine will perform collaborative filtering based on location, whereas if the user is logged in with personal account, similar content is recommended from user's past profile. Unlike previous researches, this paper focuses on the application of DL appraoch to design a

recommender engine.

The restaurant recommendation system is highly dependent on the location. Mostly, restaurant recommendations are required by travelers who do not have any prior experience of resturants in their particular visiting location. Hence, the suitable recommendation system is collaborative filtering, so that a user can be recommended with a resturant based on other

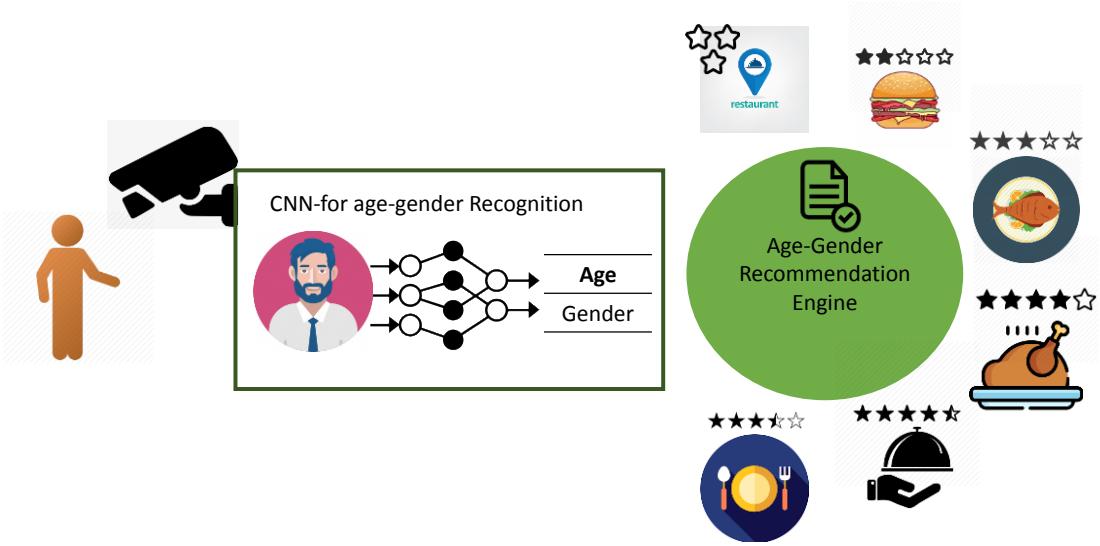


Fig. 1. Illustration of the age-gender recognition-based recommendation system.

similar users review ratings [5]. The other important requirement is the determining of the parameters for performing collaborative filtering, such as user demographics, which are highly correlated with the selection of restaurants [6]. The age and gender of person can be detected using DL and the most favourite model of all for detecting faicial features is Convulotional Neural Networks (CNN) [7].

Since this work aims at the development of recommendation system for Asians, a generic faicial recognition system may not be able to detect the proper age of the Asians [2]. To this end, age recognition system is developed as Asians have very unique facial features, which make the age and gender classification a very complex process. This work includes a CNN model that is trained using Asian dataset for detecting the age groups and gender for a user image that is given as input. Moreover, the collection of restaurants data through web crawling from Instagram is performed, which includes user rating and reviews with age groups. Based, on the collected data the pre-trained DL model recommends best restaurant to be visited by users as shown in **Fig. 1**.

The rest of the paper is organized as follows: section II explains related work. Section III focuses on system design and working. Section IV presents the details about the results and section V concludes the paper.

2. Related Work

Recommender systems provide the basis for providing best product experience to users. Hence, in our daily routine we are surrounded by several recommendation systems, e.g. writing email with Gmail, Grammerly, Youtube, Facebook, Netflix, Amazon, Maps and many more are driven by recommendation systems [1]. Hence, the imprtance of recommendation systems can not be neglected.

There are many ways to use technology for recommending user the meal they need. For instance, in [3] researchers have developed a smart table in which different sensors were installed. Based on the user data and envoirmental data, the recommendation system provides a healthy meal to its users. Basically, there are two procedures followed by recommendation systems, i.e., collaborative filtering, which recommends based on the behaviour of similar demographics [4], and

content filtering, which is based on user to user choices for a similar content [5]. The research work in [4,5] expalins the restaurant prediction using collaborative filtering and content filtering. Both the methods are valid, but highly dependent on the use case. The purpose of this paper is to recommend the best suitable restaurants to new users who do not have past history of using a particular resturant in locality. Hence, collaborative filtering is the recommended approach as explained in [6].

The top industry recommendation mechanisms include Youtube that presents the content in a way that a user is interested in and keep watching for hours [8]. Similarly, authors have explained Facebook group recommedation in [9].

The CNN is widely used for image data and researchers in [7] have explained WC-CNN model to recognise age and gender of users. A research work on an Asian dataset for age-gender recognition was achieved using Metric Regression with CNN (MR-CNN) [10]. There are many avaiable CNN models to achieve quality results and enhance the efficiency and training of models [11].

This manuscript provide a novel approach in which it combines the facial recognition with restaurant recommendation, where it first recognizes the age and gender of the user. Then based on the recognized age and gender groups, the recommendation system provides a list of the recommended restaurants in the locality.

3. System Design and Working

This manuscript proposed a DL-driven restaurant recommedation system using facial recognition. The system based on the incoming image from the camera recognizes age and gender of the user and recommends best rated restaurant in the locality. This system contributes in twofold: 1) age-gender recognition for Asian facial appearance, 2) applying collaborative filtering based on facial features to recommend restaurants.

This work designs a recommender booth that can be installed at public places like bus stops, airports, and parks. **Fig. 2**. illustartes three-step working of the system. Firstly, whenever user interacts the system, it captures the image of the person standing in front and recognises the age. Based on age and gender of the person it recommends the restaurant list with ratings. Secondly, as it is known that the restaurants always have queue to serve the customers, and customers are always uncomfortable to wait for the meal. Hence, this system provides facility of pre-ordring and prepaid recommended meal. So that, upon arrival, customer's meal can be served at the table without queuing inconvenience. Thirdly, since it is really hard to find the location of the restaurants, this system gnerates a QR code for the exact location of the restaurant, so that users can arrive at the restaurant and can have the meal with ease. The system consists of following components:

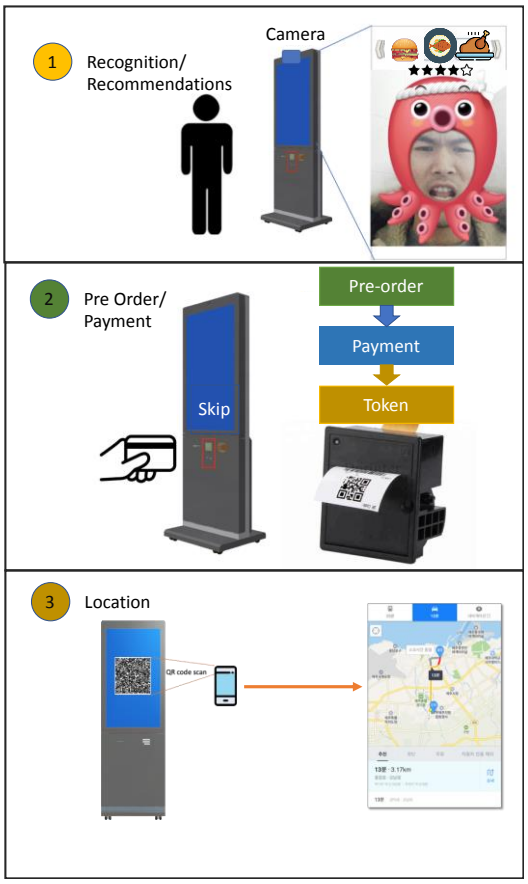


Fig. 2. Working of Age-Gender Based Restaurant Recommendation System.

3.1 Convolutional Neural Network Based age-gender Recognition:

The aim is to provide the restaurant recommendation in Korean region, so it is very important to detect the age and gender of Asian race. To this end, Asian dataset and an appropriate DL model to recognize Asian user's age and gender are required. Therefore, Asian

Face Age Dataset (AFAD) [10] was used to train the CNN model. AFAD is composed of 160 thousand images including age and gender labels. It contains the estimation of age-gender for Asian faces belonging different age groups.
Fig. 3 shows the variety of information available in AFAD dataset, where y-axis shows the number of images, x-axis depicts the age, blue bars are used for male gender, and red bars show female gender.

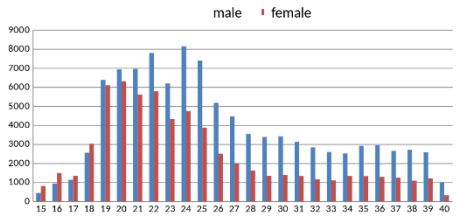


Fig. 3. Illustration of age and gender for AFAD dataset [10].

The EfficientNet: Convolutional Neural network (Convnet) is used to train the AFAD dataset. This model achieved 84.3% first-ranked accuracy on ImageNet data, and it is 8.4 times smaller and 6.1 times faster than the existing state-of-the-art CNN models [11]. Usually, the accuracy of different CNN models is enhanced using scaling of different dimensions, such as length, width, and depth of the model. The EfficientNet was a result of observation that the scaling of all the dimensions is highly dependent and proposed a compound scaling approach for enhancing the existing CNN models. It has different variants from B0 to B7 and efficiency of each of them is presented in [11], whereas compound scaling is illustrated in in **Fig. 4**.

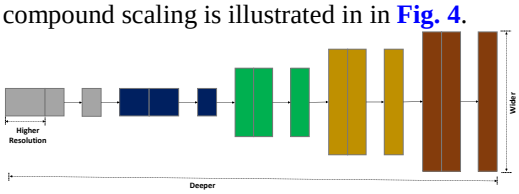


Fig. 4. Illustration of EfficientNet-based compound scaling of CNN model [11].

3.2 Recommender System

This paper uses collaborative filtering for the recommendation of restaurants based on age and gender groups. As there are not sufficient datasets available for Korean restaurants, so web-crawling over Instagram is performed to collect restaurant ratings and user reviews with their respective age and gender. The overview of the dataset is depicted in **Table 1**, where name of the restaurant with the visits of different age and gender groups is shown. The grouping is performed based on gender and age, where from 0 to13 years are placed in one group and similarly 13-19, 19-35 and 35-56 are placed in their respective groups. The dataset collected contains 1500 restaurants and ratings are calculated. Table. I. illustrates the dataset after preprocessing and classifying it into age and gender.

Table 1. Restaurant rating score data with age groups and number of visits.

restaurant_name	No	0m	13m	19m	35m	46m	56m	0f	13f	19f	35f	46f	56f	score
거명국수	38	0	1	0	0	0	0	3	0	1	0	0	0	3.1
고등어쌈밥	55	0	0	3	0	0	0	0	0	0	0	0	0	4.1
고로케야	59	0	0	0	5	0	0	0	0	0	0	0	0	4
광명식당	87	0	0	0	0	0	0	0	0	0	0	0	0	4.1
그조록	120	0	0	4	0	0	1	0	0	0	0	0	0	3.7
기억나는집	129	0	0	0	0	0	0	0	0	0	0	0	0	3.4
남경미락	161	0	0	0	0	4	0	4	0	0	0	0	0	3.7
남춘식당	165	0	0	0	0	0	0	0	0	5	0	0	0	3.9
늘봄흑돼지	188	0	0	0	0	0	5	0	0	0	0	0	0	3.4
다정이네	210	0	0	0	0	0	0	0	5	0	0	0	0	3.4
더리트리브	241	0	0	0	0	3	0	0	0	0	0	0	0	3.5
덕성원	247	0	1	2	0	0	0	0	0	0	0	0	0	3.4
덕승식당	251	0	0	0	2	0	0	0	0	0	0	0	0	3.7

The methodology used for collaborative filtering is matrix factorization (MF) and the Objective function used is single value decomposition (SVD) [12]. Python surprise library is used to execute MF on the dataset. The MF uses an embedding approach and sets a scalar value between [1, -1] (preferences, groups) for each group of users. The top row (yellow color) of the matrix represents hotel ratings for children. In the example, values near to positive 1 mean the restaurant are good for children and near to -1 mean, the restaurant is not well rated for children. After applying MF, the objective function is applied, which is SVD as represented in equation (1) [12].

$$f = \sum_{(x,y)}^{n} \left(F_{xy} - \cup x \,.\, R_y\right)^2 \qquad (1)$$

In equation (1), x and y are the observed instance values, F_{xy} is the feedback matrix, and U_x is the user embedding matrix, while Ry is the restaurant embedding matrix. Using the objective function, the feedback matrix is generated having single value either 1 or 0, as shown in **Fig. 5**. The goal of the objective function is to generate F_{xy} as close to Ux as possible.
Using the reviews and age gender, the SVD matrix was created to determine the recommended restaurants to the users.

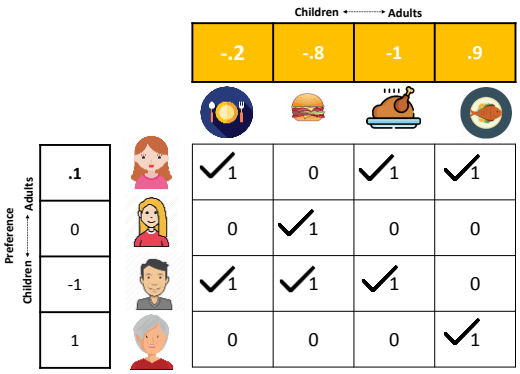


Fig. 5. Illustration of Matrix factorization-based embedding for our dataset.

4. Results

Several results need to be considered for the validation of the proposed system. The training results for the CNN are as follows: in the initial stage, the model is trained to predict the exact age and gender for the input image of the user. In this stage, the model accuracy remained close to 68 %. While testing the model it was observed that every wrong age prediction was leading/lagging by 5 to 6 years. Hence, the input age labels for images in the dataset were converted with age groups as illustrated in **Table 1**. This resulted in a significant improvement of accuracy to 81%.

Secondly, the evaluation of recommendation engine is performed based on Mean Absolute Error (MAE) value. It is calculated using the equation 2.

$$MAE = \sum_{n=1}^{n} ([a_n - b_n]) \frac{1}{N} \quad (2)$$

In equation (2), a_n is the predicted value derived from equation (1), and b_n is the actual value. Hence, this equation provides the difference between actual and predicted values. As the rating was set between 1 to 5, MAE calculates the difference between the actual value and predicted value. Hence, the averaged MAE value must be below 5 and closer to 0. The MAE value for our model is 0.27, whereas the Root Mean Square Error (RMSE) is 0.35 and 0.27.

4. Conclusions

In this paper, an AI-driven restaurant recommendation system for users based on age-gender recognition is presented. The CNN model is used to recognize the age-gender of an input image while collaborative filtering is used to recommend the best restaurants in the locality. Moreover, a real-time system that can be deployed at public places, so that the users can have access to the best recommended restaurant with convenience as well as in accordance with their age and gender. This work successfully achieves the age-gender recognition for Asian faces as well as collects its own dataset from Instagram to be used for recommendation system. In the future, the proposed approach will be extended to the global datasets and implement other approaches like content filtering for recommending restaurants to users based on their profiles.

Acknowledgements

This research was supported by the Korea Research Fellowship program funded by the Ministry of Science and ICT through the National Research Foundation of Korea (2019H1D3A1A01102980).

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2016R1D1A1B01016322).

References

[1] Oh, Y., Choi, A. & Woo, W. u-BabSang: a context-aware food recommendation system. J Supercomput 54, 61–81 (2010). https://doi.org/10.1007/s11227-009-0314-5

[2] Helen Knaggs, “Chapter 9 - Skin Aging in the Asian Population”,2009, Pages177-201, https://doi.org/10.1016/B978-0-8155-1584-5.50013-2.

[3] S. Nicolau, A. Garcia, X. Pennec, L. Soler, and N. Ayache, “An augmented reality system to guide radio-frequency tumor ablation,” *Computer Animation and Virtual World*, Vol 16, No. 1, pp. 1-10, 2005.

[4] U. Farooque, B. Khan, A. B. Junaid and A. Gupta, "Collaborative Filtering based simple restaurant recommender," 2014 International Conference on Computing for

Sustainable Global Development (INDIACom), New Delhi, 2014, pp. 495-499, doi: 10.1109/IndiaCom.2014.6828187.

[5] J. Zeng, F. Li, H. Liu, J. Wen and S. Hirokawa, "A Restaurant Recommender System Based on User Preference and Location in Mobile Environment," 2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), Kumamoto, 2016, pp. 55-60, https://doi.org/ 10.1109/IIAI-AAI.2016.126.

[6] Alif Azhar Fakhri, Z K A Baizal, and Erwin Budi Setiawan "Restaurant Recommender System Using User-Based Collaborative Filtering Approach: A Case Study at Bandung Raya Region", 2019, Journal of Physics: Conference Series, doi: 10.1088/1742-6596/1192/1/012023

[7] C. Huang, Y. Chen, R. Lin and C. -. J. Kuo, "Age/gender classification with whole-component convolutional neural networks (WC-CNN)," 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Kuala Lumpur, 2017, pp. 1282-1285, doi: 10.1109/APSIPA.2017.8282221.

[8] Paul Covington, Jay Adams and Emre Sargin "Deep Neural Networks for YouTube Recommendations", Google Mountain View, CA.

[9] Baatarjav EA., Phithakkitnukoon S., Dantu R. (2008) Group Recommendation System for Facebook. In: Meersman R., Tari Z., Herrero P. (eds) On the Move to Meaningful Internet Systems: OTM 2008 Workshops. OTM 2008. Lecture Notes in Computer Science, vol 5333. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-88875-8_41

[10] Zhenxing Niu, Mo Zhou, Xinbo Gao, Gang Hua. Ordinal Regression with a Multiple Output CNN for Age Estimation. CVPR, 2016.

[11] Mingxing Tan and Quoc V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks", 2020, eprint: 1905.11946.

[12] Dheeraj Bokde, Sheetal Girase and Debajyoti Mukhopadhyay, "Matrix Factorization Model in Collaborative Filtering Algorithms: A Survey", Procedia Computer Science, Volume 49, 2015, Pages 136-146, https://doi.org/10.1016/j.procs.2015.04.237.

Classification of knee Osteoarthritis in Anterior Cruciate Ligament with Convolution Neural Network

Mazhar Javed Awan[1,2], Mohd Shafry Mohd Rahim[1], Naomie Salim[1] and Hassan Shabir[2]
[1]School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia
Skudai, Johor Bahru, 81310, Malaysia
[2]School of System and Technology, Department of Software Engineering, University of Management and Technology, Lahore 54000, Pakistan
[e-mail: awan1982@graduate.utm.my, mazhar.awan@umt.edu.pk]
*Corresponding author: Mazhar Javed Awan

Abstract

Anterior Cruciate Ligament (ACL) is most communal injury found in human specially sportsmen. There is a need to classify the ACL tear before fully ruptured to avoid osteoarthritis. The aim of this research to classify ACL tears with deep learning model of convolutional neural network (CNN) of Magnetic Resonance Images (MRI) in an efficient manner. In this paper we trained our model on CNN of two different structures in MRI knee analysis of healthy and full ruptured ACL tears. The dataset has different types of classes that are non-injured, partial tears and fully ruptured. The experimental results showed that our proposed customized CNN model achieved 95 % accuracy that is 13% more than of standard CNN architecture on healthy and ruptured classes. The model showed good result overall and in future we can improve it to apply other CNN architectures to detect and segment other ligament parts as well.

***Keywords*:** Anterior Cruciate ligament, osteoarthritis, deep learning, classification, detection, Convolutional neural network

1. Introduction

The Anterior cruciate ligament connects the femur to the tibia and it's a very common sports medicine injury that one extensive surgery. It caused by very high forces in the ligament. Every year's there about 250,000 people that have ACL ruptured and half of these having reconstructive surgery.it is observed that 30- 60 % of those who undergo ACL surgery, develop osteoarthritis within 5 years [1]. There are various methods to identify the osteoarthritis in the ACL tear of knee by looking at the loads from gate, the biochemical changes and through radiology images like X-rays, CT scans and MRI [2].

Knee osteoarthritis (KOA) is a degenerative develops slowly over time), serious and painful that effects the large population worldwide in all age group of people Knee osteoarthritis is caused by Breakdown of cartilage and ruptured in the anterior cruciate ligament [3]. The Fig. 01 shows the damage in the anterior cruciate ligament. The ACL is one of the most commonly injured ligaments in the knee. The ligaments are strong non elastic fibers that connect our bones together. The ACL crosses inside the knee connecting thigh bone to the leg. It provides stability to the knee joint. The ACL tears can cause knee pain swelling, instability and osteoarthritis [4, 5].

Fig. 1. Ruptured ACL tear in the knee bone

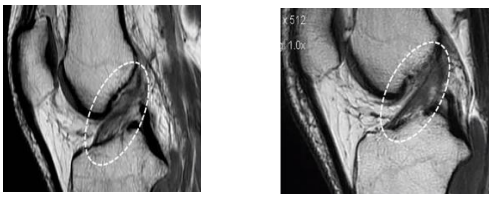
The image segmentation provides clinically important to find out the bone tissues as well as classification through MRI on the segmentation result [6, 7].

The MRI is a 3D picture that slices through the knee in three planes: sagittal, coronal and axial. The ACL goes in a 45-degree direction and lives in the middle of the knee and it prevents the tibia bone from moving forward. The PCL arches across the back of the knee and prevents the whole knee from hyper-extending. The easiest way to find the ligaments is on the sagittal, which is the side view [8]. There are two versions that are bone light version and bone dark version of the sagittal. The **Fig. 2. (a)** shows the ACL is dark and it is 45-degree direction and there is left side little PCL in the bone dark as shown in **Fig. 2. (b)**. so the bone dark type image, the PCL arching across the back so the normal, nice dark PCL.

(a) (b)

Fig. 2. Light and Dark bone of ACL and PCL MRI a) Bone Light ACL Sagittal view b) Bone dark Sagittal PCL view

There are some normal ACL with different forms of tears the ACL in **Fig. 3. (a)** has the problem of wavy type fibers. They are not real tight. They are light in colour and they are not all going in the same direction. They tend to take this wavy sort of courses. Whereas on the healed image or the improved image over **Fig. 3. (b)** There is a much more normal appearing ACL. Nice tight fibers that all go in the same direction which are dark [9].



(a) (b)

Fig. 3. Normal ACL fibers a) Normal dark fibers b) Normal tight fibers

The **Fig. 4. (a)** shows ACL which has a couple of different problems, the bottom portion clearly see there is no dark connecting to the bone in the front, in the middle and top there is no ACL no dark ACL so here completer tear. Now the PCL here **Fig. 4. (b)** Arches across the back where fibers down appear more normal but also fibers which are pretty wispy and thin showing abnormal PCL [10,11].

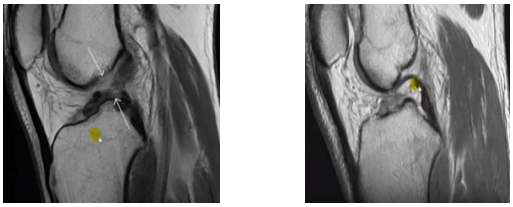


Fig. 4. ACL and PCL MRI abnormal tear a) ACL complete tear b) PCL abnormal tear

The study collected a large dataset MURA [12] related to lower extremity radiographs (X-rays) of normal or abnormal various parts of body. The CNN of 161 layers densely connected network achieved an AUC of 0.880 on this abnormality classification task. The drawbacks of this study were retrospective study with only one institution of taken binary classes, patient population and resizable images after down sampling result not good [13].

The related work [14] of the Knee MRI in which accuracy, degraded in case of partial and healthy MRI through features extraction techniques of machine learning. The study proposed CNN in

cartilage morphological degenerative changes in ACL subject specificity 85.2% only. The work was uncertain in the image annotation and was difficult to detect anomaly from infected parts [15].

Our aim is to propose such a model which can detect ruptured anterior cruciate ligament accurately which causes osteoarthritis in human.

The key contributions of this paper are as follow:

We propose an improved CNN model after adding more layers to detect osteoarthritis from MRI.

Moreover our research will help surgeons and patients.

2. Methodology

We have used dataset of images of size (330,330,32) of width, height, depth respectively where depth is the number of slices in the images. Here the problem is that image size is too big to handle on low cost and our region of interest is also a small part inside the image. So we decided to take only my region of interest and then apply deep learning model on our images. The **Fig. 5** is about region of interest extracted from original image.

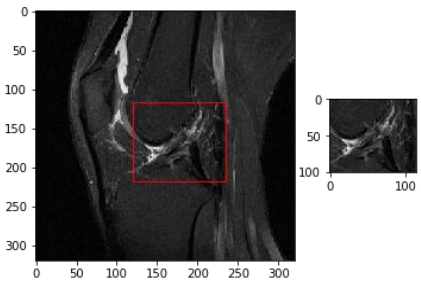


Fig. 5. Image cropped based upon interest of region

The dataset of training examples (772, 90, 90, 1) where (90,90,1) is the size of our region of interest part and testing examples test (331, 90, 90, 1). Here in the pie chart **Fig. 6** have the images of bone ruptured images in Green part with label 2, half ruptured images in Orange part with label 1 and Fine Images in blue part with label 0.

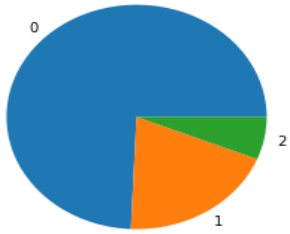


Fig. 6. Volume of three classes

2.1 Standard Convolutional Neural Network

The convoultional is feed forward neural network that is generally used to analyze visual images by processing data with grid like topology. A CNN is also known as "ConvNet".Convolutional Neural network have three features for reducing the parameters in the neural network Sparse interaction between the layers ,Parameter sharing and Equi-variant representation.

The Convolutional neural network analyses the images by extracting features from image data by applying convolutional on it. The process of applying filters makes mask on image in the form of matrices and get features from images. The objective of activation function Relu to get output from neural network and the end hidden layers are processed to get information resides in the images which can further evaluated on unseen images to get prediction of classes.

2.2 Customized CNN

Our customized model of CNN has four convolutional layers and dense layers. The standard model has over fitting so to handle this we added more dropout layers in our model. We have 1,134,455 trainable parameters. The last layer we are using softmax activation for getting probabilities of all the classes.

3. Evaluation Results

The metrics we used here the accuracy on training and test dataset. The **Fig. 7** shows the accuracy of training and test on standard CNN model after applying 100 epoch. The accuracy of test dataset is 82%.

Fig. 7. The Standard CNN model accuracy on training and test

The other graph in **Fig. 8** shows the train and test loss. After 20 epochs this model occurred overfitting.

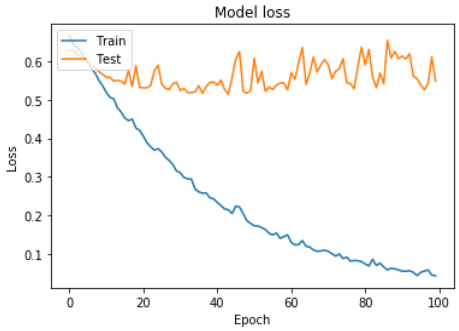


Fig. 8. The Standard CNN model loss on training and test

The **Fig. 9** shows the accuracy of our customized CNN architecture on the training and test dataset. The accuracy of test data is 95%.

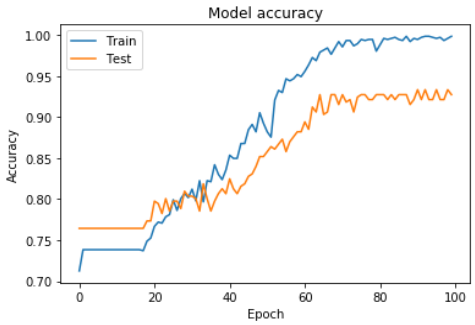


Fig. 9. The Standard CNN model accuracy on training and test

The **Fig. 10** shows the loss error in training and test data which is much lesser as compares to standard model.

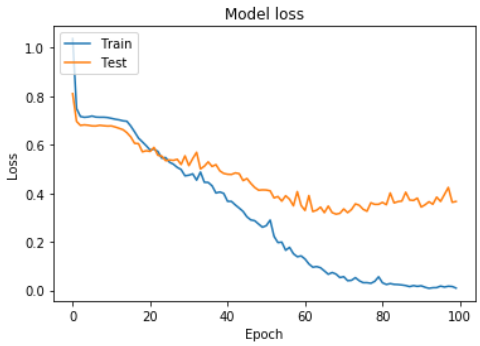


Fig. 10. The customized CNN model loss on training and test

4. Conclusions

The magnetic resonance imaging is better for predicting knee osteoarthritis. Our customized CNN model achieved 95% accuracy which is better than standard CNN model. In future we can do class balance and data augmentation to improve further.

References

[1] Spindler, Kurt P., and Rick W. Wright. "Anterior cruciate ligament tear," *New England Journal of Medicine*, Vol 20, No.359, pp. 2135-2142, 2008.

[2] Zantop, Thore, Wolf Petersen, and Freddie H. Fu. "Anatomy of the anterior cruciate ligament." *Operative techniques in orthopaedics*, Vol 15, No. 1, pp. 20-28, 2005.

[3] Naraghi, Ali M., and Lawrence M. White. "Imaging of athletic injuries of knee ligaments and menisci: sports imaging series." *Radiology*, Vol 281, No. 1 pp. 23-40, 2016.

[4] Awan Mazhar Javed, Mohd Shafry Mohd Rahim, Naomie Salim, Ajune Wanis Ismail, and Hassan Shabbir. "Acceleration of knee MRI cancellous bone classification on google colaboratory using convolutional neural network," *International Journal of Advanced Trends in Computer Science and Engineering*, Vol 8, No. 1.6 Special Issue, pp.83-

88,2019.doi:10.30534/ijatcse/2019/1381.62019

[5] LaPrade, R.F., Engebretsen, A.H., Ly, T.V., Johansen, S., Wentorf, F.A. and Engebretsen, L.,The anatomy of the medial part of the knee. *JBJS*, Vol *89*(9), pp.2000-2010,2007.

[6] Rad, A.E., Rahim, M.S.M., Kolivand, H. and Amin, I.B.M., Morphological region-based initial contour algorithm for level set methods in image segmentation. *Multimedia Tools and Applications*, *76*(2), pp.2185-2201, 2017.

[7] Norouzi, A., Rahim, M.S.M., Altameem, A., Saba, T., Rad, A.E., Rehman, A. and Uddin, M., Medical image segmentation methods, algorithms, and applications. *IETE Technical Review*, *31*(3), pp.199-213, 2014.

[8] Kam, C.K., Chee, D.W. and Peh, W.C., 2010. Magnetic resonance imaging of cruciate ligament injuries of the knee. *Canadian Association of Radiologists Journal*, *61*(2), pp.80-89.

[9] Kocabey, Yavuz, Onur Tetik, William M. Isbell, Ö. Ahmet Atay, and Darren L. Johnson. "The value of clinical examination versus magnetic resonance imaging in the diagnosis of meniscal tears and anterior cruciate ligament rupture." *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 20, no. 7 ,pp. 696-700, 2004.

[10] Grassi, A., Bailey, J.R., Signorelli, C., Carbone, G., Wakam, A.T., Lucidi, G.A. and Zaffagnini, S., 2016. Magnetic resonance imaging after anterior cruciate ligament reconstruction: a practical guide. *World journal of orthopedics*, *7*(10), p.638.

[11] Nenezić, D. and Kocijancic, I., 2013. The value of the sagittal-oblique MRI technique for injuries of the anterior cruciate ligament in the knee. *Radiology and Oncology*, *47*(1), pp.19-25.

[12] Rajpurkar, P., Irvin, J., Bagul, A., Ding, D., Duan, T., Mehta, H., Yang, B., Zhu, K., Laird, D., Ball, R.L. and Langlotz, C., 2017. Mura: Large dataset for abnormality detection in musculoskeletal radiographs. arXiv preprint arXiv:1712.06957.

[13] Varma, Maya, Mandy Lu, Rachel Gardner, Jared Dunnmon, Nishith Khandwala, Pranav Rajpurkar, Jin Long et al. "Automated abnormality detection in lower extremity radiographs using deep learning." *Nature Machine Intelligence* 1, no. 12 (2019): 578-583.

[14] Štajduhar, Ivan, Mihaela Mamula, Damir Miletić, and Gözde Ünal. "Semi-automated detection of anterior cruciate ligament injury from MRI." *Computer methods and programs in biomedicine* 140 (2017): 151-164.

[15] Pedoia V, Norman B, Mehany SN, et al. 3D convolutional neural networks for detection and severity staging of meniscus and PFJ cartilage morphological degenerative changes in osteoarthritis and anterior cruciate ligament subjects. Journal of Magnetic Resonance Imaging : JMRI. 2019 Feb;49(2):400-410.
DOI: 10.1002/jmri.26246.

[16] Chang, Peter D., Tony T. Wong, and Michael J. Rasiej. "Deep learning for detection of complete anterior cruciate ligament tear." *Journal of digital imaging* 32, no. 6 (2019): 980-986.

Real-time Vehicle Tracking Management System using GPS & GPRS

Afnan Badar Sadhayo, Dr Qasim Ali Arain, Dr Isma Farah Siddiqui, Zahid Hussain Khaskali
Department of Software Engineering, Mehran University of Engineering and Technology,
Jamshoro, Sindh, Pakistan
[e-mail: Afnan5sadhayo@gmail.com, qasim.arain@faculty.muet.edu.pk,
isma.farah@faculty.muet.edu.pk, zahidhussainkhaskheli@gmail.com]

Abstract

This paper mainly focuses on the development of an app that improves the problems of transportation of buses. This is an android application. The main aim of this app is to try to reduce the waiting time of travelers at bus stations by sharing the bus details between drivers and travelers. This app gives the exact time of arrival and departure of buses so the travelers so that they can easily wait. The travelers will be informed by the most important feature provided by our app which is Real-time location tracking of Bus & ETA will be shown on check-post about Bus also SMS alerts and email notifications in the case when the buses are near to travelers nearest stations. travelers can easily update themselves about the bus's exact locations and the routes by just looking at the electronic map already installed in our app and can be used via the internet they can see the ETA of the bus, at what time the bus will arrive on the selected Check-post. This app is designed in such a way that it gives travelers real-time services and will be very useful for them. The case study is been done for Mehran University Students Bus and the app is targeting the students and routes within the MUET campus.

Keywords: Vehicle Tracking System, GPS, GPRS

1. INTRODUCTION

The principles of information technology are growing very quickly and any job or order can be placed from home with mobile applications or websites. Due to the relentless growth in automobile applications, efficient transportation has set off a very significant feature of today's era. To support their students, learning institutions provide transportation facilities. Various attempts have been attempted to enforce transport protocols, and further, consistently to reduce student difficulties with their concerns, i.e. due to the uncertainty in scheduling, the student delays their bus, which is a system malfunction or drivers do not follow the route. Vehicle monitoring is one of how this dilemma is minimized by letting students know what is happening in real-time. A web server uses the orientation of a car to relay the position message accurately to the user. This research presents a vehicle tracking system distributed in three modules Global Positioning System (GPS) [1] for positioning, Web server (GPRS) [2] for data storage and transmission, and Client i.e. Student where Google maps software in incorporated for displaying the location. For the Intelligent Transport System (ITS) [3] the tracking system can be very useful. In probe vehicles, it can, for instance, be used to assess the congestion area in real-time traffic data. Whenever an incident happens, in case of emergency it can quickly report the vehicle position to its nearest rescue agent.

Our proposed system is the Hybrid mobile app which is used to discover the location of MUET buses using GPS. This app enables us to compute the easy and precise location of points

(buses) & provide the student with all necessary details regarding the MUET Points (buses), likely the arrival time of points, route of the points, its exact location, expected to wait time for the point. SMS alerts and notification alerts when the point (bus) is near to your check post and expected time to wait for the bus. All these pieces of stuff can be viewed on digital mapping with the help of our app via the internet. We call our real-time tracking management system as MUET POINT TRACKER.

The fundamental goal of this application to solve the difficulties of the MUET students when it comes to coaches is to cut students' waiting time at bus stations. By using this app, you have no trouble with buses such as how long you have to wait for busses when you arrive at the bus and where the bus stands. Provide all coaches with their coach numbers with the location of a bus heading to the stop and give this information to students on the destination you want to meet. The architecture of the client-server. This app is not expensive, since all students at Mehran University need their ID and password, as it is important to use the features of this app such as the location of buses on their scheduled time of arrival on the bus list and notify students that the bus is its closest stop and has several additional features.

1.1 PROBLEM STATEMENT

In this present world time is very precious for all of us especially for students. Being a product of modern technology, cell phones are widely used and have become more popular. In order to facilitate their students, academic institutes provide transport services. Daily routine of student's vehicle/Bus transportation system faces a number of different unknown challenges as the day progresses, for instance:

- Traffic congestion
- Unexpected/Unknown delays
- Irregular Points-dispatching times
- Other incidents.

This leads to students frequently having to wait at their bus/position stops for some time because they do not know when the bus/point arrives. The effect is that students are often late for classes. So, the time of arrival of the bus/bus point cannot be determined, so instead of using alternative transport, they simply want to wait for their respective point.

1.2 OBJECTIVES AND GOALS

The main aim of this project is to provide the point students with guidance/assistance and help them identify the points and observe the time required to arrive from the point to their nearest stop. The key objectives of this framework are:

- Low hardware cost with the power performance.
- Low power consumption.
- Provides security of data privacy as the application is hosted on servers centered on the university.
- Procuring a GPS device to use in points through which the application will show the location of the point.
- To make more accessible and constantly updated dynamic Points information to Students.
- Provides information about all the MUET Points along with the Points number, Driver name & details that go through the route.
- Real-Time tracking the position of the MUET Points on digital maps via an online monitoring program.
- Acknowledging the position via GSM, SMS.

1.3 EXPECTED OUTCOMES

The expected result of our project is possibly a fully functional "MUET POINT TRACKER" website application that will enable our users to locate their busses on Google Maps in a real-life location and to receive message notices about their university busses. Installed on android and IOS smartphones, this website app allows an Internet connection to obtain those locations and updates. You shall be told that your bus is placed so that your point is not missed and your bus is not missed and your losses are not removed.

1.4 LIMITATIONS

Both licenses of used components must also be checked in order to be able to go live with the device.

2. LITERATURE REVIEW

This chapter covers the contextual information and related study for this research. The main objective of this chapter is to highlight the working of Vehicle Tracking System. We cannot ignore the problems which have been faced by our students of MUET regarding buses. Mostly they waited for the buses at the bus stations without knowing when the bus is arrived which is so irritating and illogical. Student don't know where the buses are what are their exact locations and when they will arrive and what time does it require on arrival so the students can easily wait. We know that in our university there is no fix time of closing the departments every department close according to its time because of time variations and every batch has its own schedule of classes. So students must know the current location of buses so they can wait instead of walking the whole campuses for finding where the bus is. And at the morning the points time is fix but sometimes the arrival of buses is before their fix time and sometimes after which creates problems for most students.

2.1 FIELD OF THE INVENTION

The current invention presented has its roots in a communication network and specifically to a system and processes for detecting vehicle information. Nowadays, people are communicating via wireless networks which transmit signals for sending messages from one node to another node. As a matter of fact, one cannot determine the location of an object via these wireless networks. For that, we have Global Positioning System (GPS) fixed in a device in order to direct an object which is not an expensive alternative. Therefore, it is desirable to have a network in which users from different networks can communicate with one common node for providing its information. Thus, it can identify the location of any remote object by an alarm.

U.S. Pat. No. 20100094482 A1 describes a communication system for tracking and controlling vehicles via wireless communication signals. The vehicle tracking system proposed in this patent comprises a vehicle tag and locator device for sending and receiving wireless signals and determining the relative position of the vehicle respectively [1]

U.S. Pat. No. 5218367 A describes a way of tracking vehicles movement used in an emergency. The components of this tracking system are cellular processing unit for installation in a vehicle, alarm sensors associated with the cellular unit, the sensors, emergency switch, one cellular antenna and a remote alarm monitoring station together with a computer associated to a cellular processing unit. [3]

U.S. Pat No. 74533356 B2 describes a vehicle tracking system for tracking motor vehicles for law enforcement purposes. It comes with a device that can be magnetically and mechanically attached to a vehicle, the device comprises a GPS receiver and a data transmitter. [2] The GPS receiver delivers to the data transmitter the latitudinal and longitudinal coordinates when the device is initiated. This presents the current location of the vehicle. It's one way of preventing high-speed chases done by law enforcers because due to these chases a big amount of collateral damage takes place. This system helps in tracking those particular vehicles that are subject to a police pursuit. Using this, the law enforcers can monitor and track their prey.

U.S. Pat. No. 5515043 A describes the monitoring mechanism for the vehicle that enables the user to attach remotely to the vehicle and activates an alarm system. In the case of misfortune, emergency, or accidents, it may also automatically telephone one or more pre-programmed numbers. Upon remote access by telephone or PC, when a valid PIN is given, the user can obtain position, speed, and command menu options to trigger an alarm. [4]

U.S. Pat. No. 6873912 B2 describes a tracking system for a vehicle that tracks a vehicle reliably using a matching template. "The vehicle tracking system is ideally a previous vehicle tracking system, which tracks previous vehicles." [5]

ISSN No.2348-4845 describes the usage of fleet tracking system used for localization of automobiles with the help of GSM and GPS-SMS services. The tracking system identifies

the coordinates of the automobile and sends information about that automobile to the user having a mobile phone by an SMS. The system consists of an alarm system as well. GPS receiver, microcontroller, and GSM modem are the three components of the device presented in the paper. The GPS receiver is used for satellite information through longitudinal and latitude coordinates. The GPS receiver is used. The microprocessor then manipulates the information and transfers this set of information to the user. The microprocessor is also used to get the speed of the vehicle and send it to the user. And the information which is to be transferred to the user is sent through GSM-SMS.

2.2 RELAVENT WORK

There were many no research and implementation of a vehicle tracking system using GPS and GSM services, but due to the uplift of the selective availability, the accuracy of GPS was still challenged. The accuracy of the GPS could be about 5m in heavily populated areas with large buildings always reflecting and blocking the key satellite signals. For cellular telephone positioning either a network or a mobile station, there are many technologies available. They used attenuation of the signal, the angle of arrival, the time of arrival, and the time of arrival to find the network phone unchanged. In TDMA (Multiple Access Time Division), the time of the signal to access cell phones can be used for location calculations. The GPS and GSM positioning technology is therefore sufficiently advanced for positioning precision and civil availability.

The system requires the administrator to add the driver id and password to the system. The role of the administrator is to register a new student with parent information. This project was implemented on the web for Android. The application shows the maps and routes to various locations automatically with the GPS system. The bus location will also be monitored by customer servers and sent to the customer's computer. It uses simple distance measurements for two locations and offers required information on both routes for people to pick up buses conveniently on the designated route or on any other transport possible. [6]

In order to track school buses in real-time, the app utilizes Google APIs and GPS. The application also helps to monitor pupils, parent/control contacts, and emergency alerts at periods such as pneumatic puncture, motor problems, and even injuries. For implementation, no additional hardware is required. Most functions operate in the background, so there is no need to specifically remind warnings. A method to monitor school busses and to provide more effective and efficient interaction leading to greater reliability and safety is defined by the proposed system. [7]

3. RESEARCH METHODOLOGY

3.1 SYSTEM ARCHITECTURE

It displays the system's architecture. We want to concentrate on architecture on the entities of the system and their interrelationship. It also displays the organization of the device entities. The diagram below shows our application architecture.

The methodology consists of system architecture and overall working flow of system is explained in steps.

Step 1: User login to Tracker application via authentic credentials.

Step 2: User request for the real-time location of bus/vehicle.

Step 3: Server locates the Bus/vehicle location from hardware tracker using GPS/GPRS and send to user.

Step 4: Set arriving notification of bus/vehicle at user’s marked check-post.

Step 5: Server sends push notification to application and SMS to registered user.

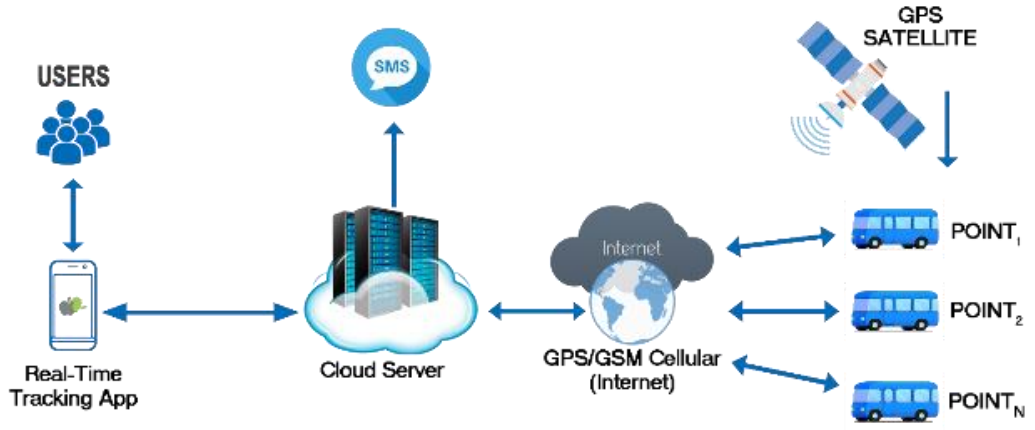


Fig. 3.1. System Architecture Diagram

3.2 SEQUENCE DIAGRAM

To visualize the partnership series of objects, a sequence diagram is used. The sequence of operations carried out in the system is defined. It describes the time-related interaction between objects. The photo. 3.2 for our program is the sequence diagram.

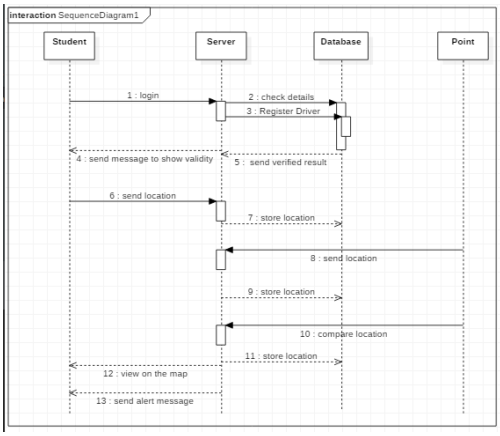


Fig. 3.2. Sequence Diagram

4. RESULT AND GRAPHS

4.1 Application Interfaces

This section covers all the interfaces designed for a client in order to make the application user-friendly. **Fig. 4.1** to **4.5** show different interfaces of developed application.

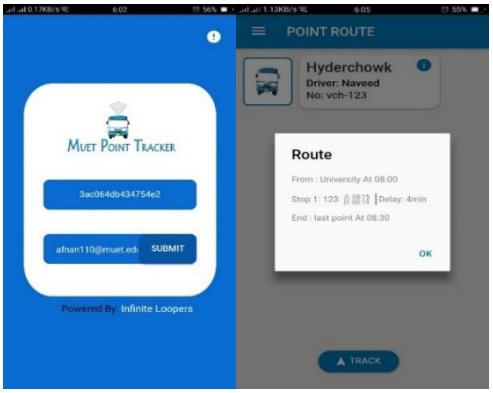


Fig. 4.1. Login **Fig. 4.2.** Routes

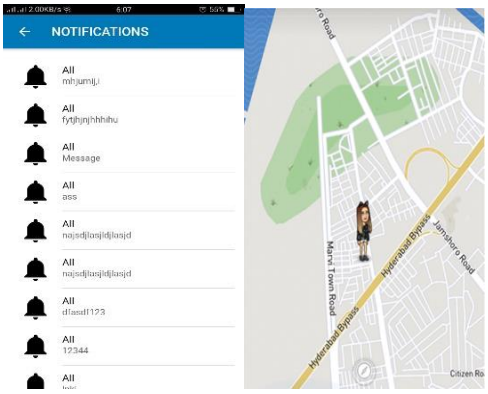


Fig. 4.3. Notifications **Fig. 4.4.** Map

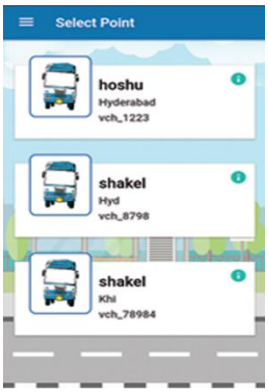


Fig. 4.5. Bus / Vehicle

4.1.2 SERVER INTERFACES

The server interface is shown in **Fig. 4.6**.

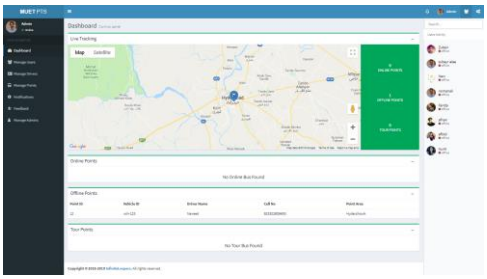


Fig. 4.6. Server Side: Dashboard of admin

4.2 SURVEY

We conducted the survey to get the response from the students of Mehran University of engineering and technology to predict the usage and need of the application. It was done through a questionnaire in which students selected the best answer the questions or multiple options if needed. The results of the survey are as follows in **Fig. 4.7**:

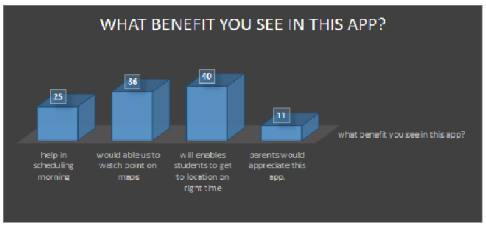


Fig. 4.7. Survey Graph-1

This histogram in **Fig. 4.8** shows the frequency of the student's response regarding the application's benefit. FOUR options were present in the questionnaire that were:

- Help in scheduling the morning.
- Would able student to watch point on maps.
- Will able students to get to the location on right time.
- Parents would appreciate this application.

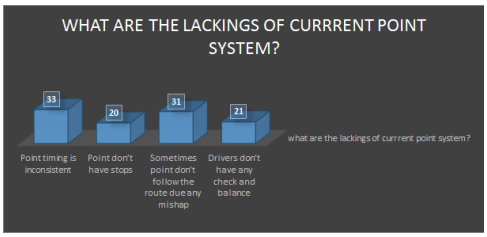


Fig. 4.8. Survey Graph-2

Before developing the application, we needed to survey the lacking's in the current transport system. So, we included the question for the lacking's of the current system and results were found as above in the graph **Fig. 4.8**. This question also had four options from which participants had to select one or multiple, which were:

- Point timing is inconsistent.
- Point don't have stops.
- Sometimes point doesn't follow the route due to any mishap.
- Drivers don't have any check and balance.

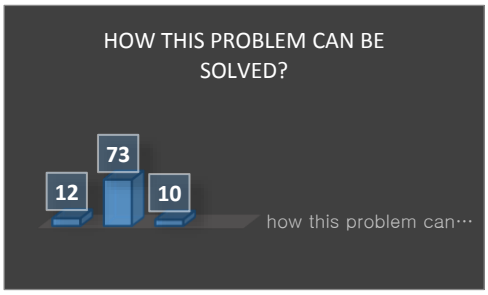


Fig. 4.9. Survey Graph-3

After getting the statistics about the current problems that students face in their transport, we placed another question that how to solve these problems. As it is an application developed for the university students so their view is important as a requirement. The students selected their choice from the following options.

- By getting to the stop early.
- By having this application.
- Having contact with the driver.

4.3 Comparison of GPS/GPRS Modules

Product	SIM7060R	A7670SA	SIM7920G-M2	SIM5320J	SIM808
Technology	LPWA	4G	LTE-A	3G	2G
General Feature	Cat-NB B1/B2/B3/B4/B5/ B8/B12/B13/ B17/B18/B19/B20/B25/B26/B28/ B31/B66/B70	LTE-FDD B1/B2/B3/B4/B5/ B7/B8/B28/B66 GSM 850/900/1800/1900MHz	LTE-TDD/LTE-FDD/WCDMA GNSS (GPS;GLONASS and BeiDou)	Dual-Band UMTS/HSDPA 850(800)/2100MHz Quad-Band GSM/GPRS/EDGE 850/900/1800/1900MHz	Quad-band 850/900/1800/ 1900MHz GPRS multi-slot class 12/10
Power Consumption	PSM: 3.4uA Sleep: 0.4mA Idle: 5.6mA eDRX@655.36s: 30uA	LTE: TBD GSM,BS_PA_MFRMS=2: TBD	3.7V	3.7v	3.7v
Interfaces	UART, SIM, SPI, I2C, GPIO, ADC	SIM Card:1.8V/3.0V	USB, UART, SIMCard, Antenna x5	USB, UART, SIMCard, keypad	68 SMT pads Analog audio interface Serial, USB, GSM/GPS antenna
Data	Up: 62.5Kbps Down: 26.15Kbps	LTE (Mbps): up: 5- Down: 10 GPRS/EDGE (Kbps) Up: 236.8 -Down: 236.8	Up: 150Mbps Down: 2.0GMbps	Up: 384Kbps Down:384Kbps	Up85.6kbps Down:12kbps
Footprint	24*24*2.6mm	26*26*2.8mm	42.0*30.0*2.3mm	30 * 30 * 2.9 mm	24*24*2.6mm
Voltage	2.1V ~ 3.6V	3.4V~ 4.2V	3.135V~ 4.4V	3.4~ 4.2V	3.4 ~ 4.4V
Temperature	-40°C ~ +85°C	-40°C ~ +85°C	-30°C to +70°C	-30°C to +80°C	-40°C ~85°C
Other feature	TCP/UDP/HTTP/ HTTPS/TLS/DTLS/ DNS/PING/LWM2M/COAP/MQT	TCP/IP/IPV4/IPV6/ Multi-PDP/FTP/FTPS /HTTP/HTTPS/DNS	TCP/IP/IPV4/IPV6 /Multi-PDP/FTP/FTPS /HTTP/HTTPS/DNS	FTP/FTPS/HTTP/HTTPS/SMTP/DNS	UMX/TCP/UDP /FTP/HTTPMMS/SMt/PSSL/DTMF
Control	AT Commands	AT Commands	AT Commands	AT Commands	AT Commands

Fig. 4.10. Comparison Chart

5. CONCLUSION

This project consists of a hybrid framework and several other features that would be useful to the system were added to the project. In order to pass from source to destination, the application offers the method of approach. A highly accurate system is implemented that offers a fantastic experience, while the student is waiting for the exact location of his/her destination. The application is highly targeted especially for those people who are not capable of leaving their private cars and reaching their desired destination in due course.

6. FUTURE WORK

The project can be fitted with a Remote/Wireless Security Camera. Parents or users can track their children on the bus using that camera, and we also can track the accident with the help of high sensitivity vibration sensors. In case of an accident on the lane, we can spot the accident using a vibration sensor and we can send it to the owner, hospital, and police. Increased precision and the most effective route found. A GPS module supporting 5G CPE, MiFi, auto, router, and other WLAN terminals should use the SIMW80 module instead. This is combined with the WLAN/BT processor, BT baseband,

and 2.4G/5G RF transceivers that increase the location accuracy of the high point. This module filters unnecessary noise and enhances the accuracy of the signal in question.

REFERENCES

[1] Reineman, B.D., Lenain, L. and Melville, W.K., 2016. The use of ship-launched fixed-wing UAVs for measuring the marine atmospheric boundary layer and ocean surface processes. Journal of Atmospheric and Oceanic Technology, 33(9), pp.2029-2052.

[2] B. Hnich, F. R. Al-Osaimi, A. Sasmaz, Ö. Sayın, A. Lamine and M. Alotaibi, "Smart online vehicle tracking system for security applications," 2016 IEEE International Conference on Big Data (Big Data), Washington, DC, 2016, pp. 1724-1733, doi: 10.1109/BigData.2016.7840787.

[3] Sheffer, E.A. and Thompson, M.J., Trackmobile LLC, 1993. Vehicle tracking system. U.S. Patent 5,218,367.

[4] Berard, A.J., Mentzer, J.L. and Nixon, D.C., Berard, Alfredo J., Mentzer, James L., Nixon and David C., 1996. Cellular/GPS system for vehicle tracking. U.S. Patent 5,515,043.

[5] Shimomura, N., Nakamura, S., Gotoh, T., Fujimoto, K. and Muro, H., 2003. A method of tracking a forward vehicle using a scanning laser radar and a camera. ITEIS, 123(8), pp.1427-1438.

[6] J. Research, "School Bus Tracking Application | J4RV4I11011", *Academia.edu*, 2020. [Online]. Available: https://www.academia.edu/40695006/School_Bus_Tracking_Application_J4RV4I11011. [Accessed: 17- Nov- 2020].

[7] ttps://www.irjet.net/archives/V4/i4/IRJET-V4I4627.pdf. [Accessed: 17- Nov- 2020].

Solving Complex Job Shop Scheduling Problem Using Self-Supervised LSTM

Xiaorui Shao[1], Chang-Soo Kim[1], Dae Geun Kim[2] and Ilkyeun Ra[3]
[1]Information system, Pukyong National University
Busan, South Korea
[2] Korea Dyeing & Finishing Technology Institute, Busan, South Korea
[3] Department of Computer Science and Technology, University of Colorado Denver, Denver, CO 80203, USA
[e-mail: shaoxiaorui@pukyong.ac.kr, cskim@pknu.ac.kr, dgkim0119@dyetec.or.kr, ilkyeun@ucdenver.edu]
*Corresponding author: Chang-Soo Kim

Abstract

The job shop scheduling problem (JSSP) is critical for smart manufacturing regarding materials arranging and resource supply. Conventional methods solve JSSP by using one optimal algorithm, which is very time-consumption and cannot deal with complex JSSP instances. This paper presents a novel deep learning-based method named self-supervised long-short term memory (SS-LSTM) to accurately handle complex JSSP. First, using the optimal method to generate sufficient training samples in small-scale JSSP. Then, SS-LSTM is applied to extract rich feature representations from generated training samples and to decide the next action. In the proposed SS-LSTM, two channels are employed to reflect the full production statues. Specifically, the detailed-level channel records 18 detailed product information while the system-level channel reflects the type of whole system states identified by the k-means algorithm. Moreover, adopting a self-supervised mechanism to keep high feature extraction capacity simultaneously ensuring the reliable feature representative ability, which is implemented by LSTM autoencoder. The authors designed, trained, and compared the proposed method with the current most accurate JSSP solver on some famous, complicated JSSP instances. The experimental analysis confirmed the effectiveness and priority of the proposed method for JSSP.

***Keywords*:** Job shop scheduling problem (JSSP), smart factory, LSTM, self-supervised learning.

1. Introduction

The job shop scheduling problem (JSSP) plays a vital role in the processing of building a smart factory regarding intelligent manufacturing [1], resource supplying [2][3], and cost-saving [4]. The traditional methods such as genic algorithm (GA) [5], particle swarm optimization (PSO) [4] aims at finding the best optimal solution for JSSP, which is very time-consumption. Moreover, it cannot solve complex JSSP as its tremendous solution space will cause NP-hard. Significantly, one $m \times n$ JSSP instance's solution will reach to $(n!)^m$. where m is the number of machines, and n is the number of jobs [6][7].

This research was supported by the Technology Innovation Program 20004205 (the development of smart collaboration manufacturing innovation service platform in the textile industry by producer-buyer) funded by MOTIE, Korea.

Some learning-based methods have been proposed to overcome the shortcomings of conventional methods. Chiefly, the deep learning-based method extracts massive attention due to its potent feature extractions capacity [8]. Currently, the most accurate deep learning-based method is developed by Zang et al. [7]. They applied a convolutional neural network (CNN) to extract feature representations to forecast the next action. However, their model requires much time to train, and it ignores some critical features with time steps because JSSP is an ongoing process-making problem. Therefore, the accuracy is still can be improved. Motivated by this, the authors proposed a novel LSTM-based method to extract rich feature representation for the next decision considering the time step's influence. In the proposed method, dual channels are employed to reflect the production environment fully. The self-supervised mechanism makes a trade-off between feature extraction and feature reconstruction. Consequently, the extracted features are accurate and robust, which enhance forecasting accuracy.

2. The Proposed SS-LSTM for JSSP

2.1 JSSP

Assume that J jobs need to be processed in M machines, each machine M_m has a particular function, and each job J_j contains a series of subphases $\{J_{j1}, J_{j2}, \dots, J_{jm}, \dots, J_{jM}\}$. Each subphase J_{jm} must be processed by corresponding machine M_m once with time T_{jm}. The whole process will end until all jobs are completed. The objective of JSSP is to find the minimum make-span. One 3×4 JSSP instance is given to explain the JSSP, as shown in **Fig. 1**. As marked with red color, the first phase of the first job J_{11} needs to be processed at machine 1 with T_{11} (10) units.

Job order

J_{1m}	1	2	3	4
J_{2m}	4	3	1	2
J_{3m}	1	2	4	6

Job time

T_{1m}	10	12	32	14
T_{2m}	40	13	11	12
T_{3m}	16	12	42	16

Fig. 1. 3×4 JSSP instance, left is the job order table while the right is the corresponding requiring time table.

2.2 Solving JSSP using Deep Learning

This paper treats JSSP as a multiple subclassification problem using deep learning technology. The workflow, including six steps, as shown in **Fig. 2**. Step 1, 2 is to obtain the training samples; Step 3 is the key to ensure the accuracy of solving JSSP; Step 4, 5, and 6 are to solve one upcoming JSSP instance.

Furthermore, using algorithm 1 generates the corresponding labels in step 2 within each machine's priority for each job. By using algorithm 1, one $J \times M$ JSSP instance would be converted into an M-classification problem.

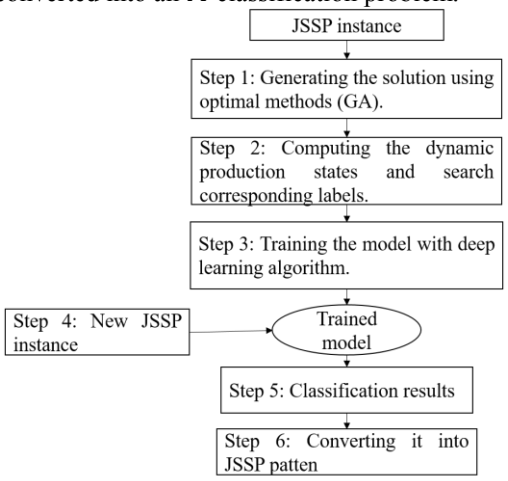


Fig. 2. Deep learning for JSSP

Algorithm 1: Labels generation algorithm

Input: Optimal solution matrix X, which records $J \times M$ elements of $\{Job_{id}, Job_{phase}\}$, where Job_{phase} is the order in the job.
Output: Sub-classification labels label.

```
for i= 1, 2, 3, …, M do:
    for j=1, 2, 3, …, J do:
        if Job_id == X[i, j, 0] and Job_phase == X[i, j, 1]
            label.append(j)
return label
```

2.3 The Proposed SS-LSTM

The proposed SS-LSTM for solving JSSP using dual channels: detailed-level and system-level channels, as shown in **Fig. 3**. Detailed-level channel records some detailed production states while the system-level channel is for system-level states. Furthermore, it consists of five steps: constructing the input sates, detailed-level feature extraction, system-level feature construction, feature fusion and self-supervised learning, output the targe and update the network, respectively.

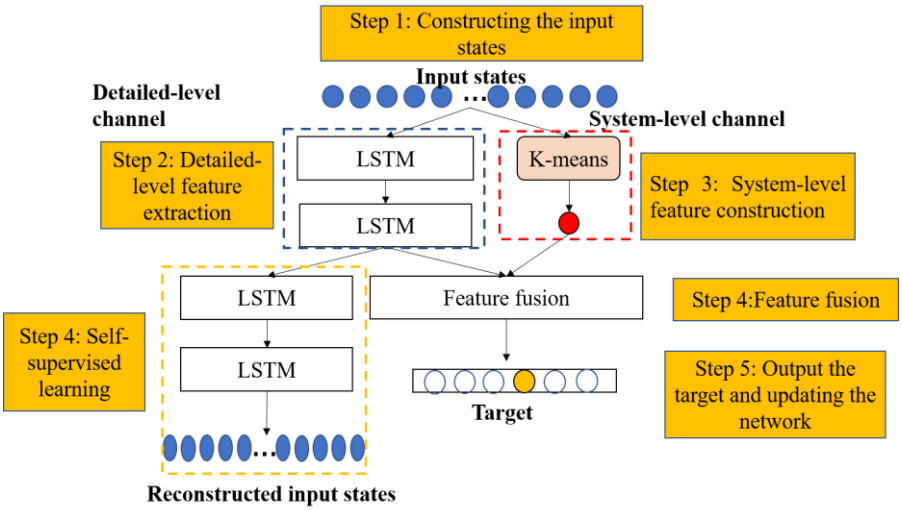


Fig. 3. The proposed SS-LSTM structure for solving JSSP

A. Constructing the Input States

The proposed SS-LSTM utilizes 18 dynamic production states to reflect the detailed production environment, as shown in **Table 1**. The input is formalized as:

$$\text{Input} = [\frac{J_{phase}}{J}, \frac{k}{n}, \dots, O_{i,j}] \quad (1)$$

B. Detailed-level Feature Extraction

The detailed-level channel utilizes two stacked LSTM layer to extract rich feature representations considering the impact on time step, as shown in (2). Where LSTM() is LSTM operation.

$$\text{Features}_{detail} = \text{LSTM}(\text{LSTM}(\text{Input})) \quad (2)$$

C. System-level Feature Construction

The previous step has extracted rich detailed-level features. The proposed method adopts the Kmeans algorithm to identify the system-level feature representations to enhance the feature representative capability again, as defined follows:

$$\text{Features}_{system} = \text{Kmeans}(\text{Input}) \quad (3)$$

D. Feature Fusion and Self-supervised Learning

To get the fusion features of detailed-level features and system-level features, this manuscript adopts one concatenate layer to combine them, as defined follows:

$$\text{Features} = \text{Concatenate}(\text{Features}_{detail}, \text{Features}_{system}) \quad (4)$$

Table 1. The defined 18 detailed-level states

The ratio of phase order $\frac{J_{phase}}{J}$, where J is the number of jobs; Position order $\frac{k}{n}$. Order of procedure in machine m; The ratio of the procedure to all machine M: $\frac{m}{M}$.
The ratio of processing time $\frac{T_{im}}{sum(T)}, \frac{T_{im}}{max(T)}, \frac{T_{im}}{min(T)}, \frac{T_{im}}{mean(T)}$
The ratio of processed J_i's processing time $T_{i,:J}$ to the total processing time of job T_i: $\frac{T_{i,:J}}{T_i}$.
The ratio of processing time T_{ij} to processed J_i's processing time $T_{i,:J}$: $\frac{T_{i,j}}{T_{i,:J}}$.
The ratio of processing time T_{ij} to total J_i's processing time T_i: $\frac{T_{i,j}}{T_i}$.
The ratio of processing time T_{ij} to unfinished J_i's processing time $T_{i,j:}$: $\frac{T_{i,j}}{T_{i,j:}}$.
The ratio of job index j to job number J: $\frac{j}{J}$. The ratio of machine m to machine number M: $\frac{m}{M}$.
The ratio of processing time T_{ij} to machine m's processing time $T_{:,m}$: $\frac{T_{i,j}}{T_{:,m}}$.
The ratio of J_i's processing time $T_{i,:}$ to machine m's processing time $T_{:,m}$: $\frac{T_{i,:}}{T_{:,m}}$.

Processing time $T_{i,j}$; Processing order $O_{i,j}$.

Meanwhile, the proposed method adopts self-supervised LSTM to keep high feature extraction capacity simultaneously ensuring the reliable feature representative ability. Two symmetrical LSTM layers are utilized to reconstruct the input states, which is formalized as:

$$\text{Input}_{\text{rec}} = \text{LSTM}(\text{LSTM}(\text{Features}_{\text{detaile}})) \quad (5)$$

E. Output and Network Updating

The proposed SS-LSTM has two outputs: reconstructed input states as described in (5) and target machine priority as follows:

$$\text{Target} = \text{dense}(\text{Features}) \quad (6)$$

where one dense layer with M nodes is employed to predict machine priority, the highest will be selected as the predicting label.

The proposed method calculates two loss functions to update the parameters of hidden layers in the SS-LSTM. One is mean square error (MSE) for reconstructing input states. Another one is categorical cross-entropy (CCE) for outputting target. The proposed method employs α, β to combine them as follows:

$$\text{loss} = \alpha\text{MSE} + \beta\text{CCE} \quad (7)$$

where $\alpha + \beta = 1$. By minimizing the loss, the hidden layers of SS-LSTM are updated. Moreover, the influence of α, β, will be discussed in the next section.

4. Experimental Verification

The authors implement the proposed SS-LSTM based on the operating system of ubuntu 16.04.3 with 23GB memory at a speed rate of 3.6 GHz, TensorFlow backend Keras.

4.1 Modeling

Configuration: This paper adopted "adam" as an optimizer to find the best convergence path. Rectified linear unit (ReLu) is selected as the activation function except for two outputs are "linear" and "softmax." The detailed configurations using an example of ten machines, as shown in Fig. 4.

Data generation: The authors randomly generate 246 10×10 JSSP instances whose processing time ranges from 10 to 100. Utilizing GA to solve those instances and calculate its detailed-level input states as defined in Table 1. Moreover, algorithm 1 is applied to create corresponding labels. At last, it obtained 24600 samples.

Training the model: This paper adopted an early stop strategy to train the model. Specifically, splitting the generated samples into two parts: 80% training samples to train the model; The rest 20% validation samples is for finding the best model within given 150 epochs using the early stop strategy.

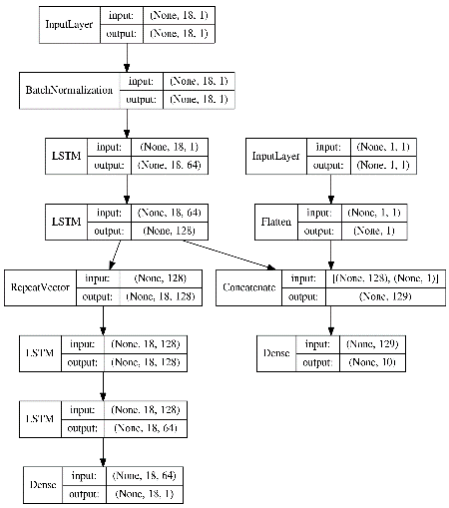


Fig. 4. The structure and detailed configuration of the proposed SS-LSTM for JSSP

System-level states identification: The proposed method adopts the K-means algorithm to identify system-level production states. This manuscript specifies cluster numbers to be 1-30. Fig. 5 shows the values of the within-cluster sum of errors (WCSS) [9] on the training set. It is clear to select 13 as cluster numbers due to changing in WCSS begins to level off at cluster 13.

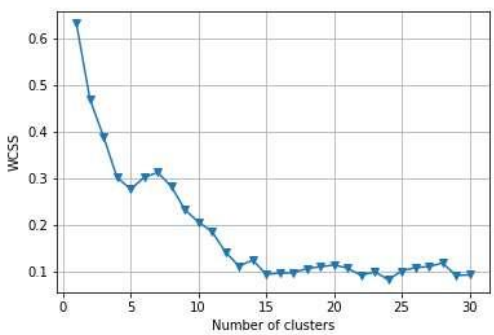


Fig. 5. The WCSS values of different clusters

4.2 The influence of α, β

After getting the two-channel inputs, the authors train the model with different α, β on one famous JSSP data set-ft10 [10] to find the best trade-off between them. The results indicate that α 0.4 and β 0.6 performs best with the make-span of 1613, which could be found in **Table 2**. Moreover, the smaller α is better, as conducted from **Fig. 6**. The sequential analysis is based on α 0.4 and β 0.6 as they perform the best.

Table 2. The results on ft10 using different α, β

α (MSE)	β (CCE)	Make-span
0.1	0.9	1853
0.2	0.8	1788
0.3	0.7	1816
0.4	0.6	1613
0.5	0.5	2037
0.6	0.4	1832
0.7	0.3	1932
0.8	0.2	1901
0.9	0.1	1896

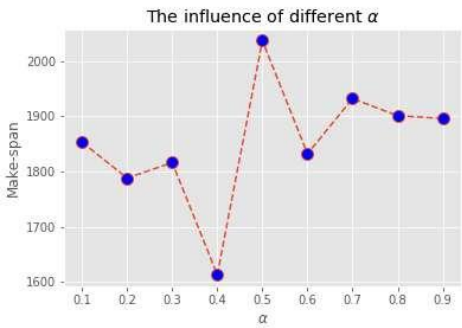


Fig. 6. The influence of different α on ft10

4.3 Comparative Analysis

To validate the effectiveness of the proposed SS-LSTM for JSSP. The authors compared it with the current most accurate method [7] on some complex JSSP instances, including ft10, ft20 [10], and la [11]. The comparison results using make-span, as shown in **Fig. 7**. The findings indicate that the proposed method wins three times over four JSSP instances except for reference [7] performs a little better on la36. Moreover, the proposed method does not require two-dimensional transformation, while reference [7] needs. Thereby it saves lots of computing resources.

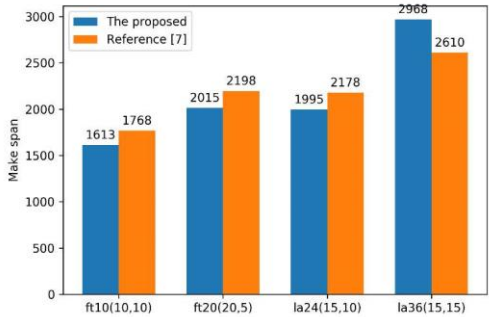


Fig. 7. The comparison results for complex JSSP. ft10 (10,10) stands by one 10 by 10 JSSP instance.

4.4 Ablation Study

To explore each component's influence on the proposed SS-LSTM for solving complex JSSP, the authors designed LSTM alone to validate the system-level channel's effectiveness. Designed SS-LSTM without a self-supervised mechanism to validate its impact. All configurations are the same as SS-LSTM, and the results are tested on the ft10 instance, as shown in **Table 3**. Comparing LSTM alone with SS-LSTM without a self-supervised mechanism, the system-level channel has reduced the make-span of 8.38%. Therefore, it confirmed the effectiveness of the system-level channel in the proposed method. Moreover, the supervised mechanism application has improved 11.18% performance by comparing it with the proposed SS-LSTM. The findings have confirmed the effectiveness of each part in the proposed SS-LSTM.

Table 3. The ablation analysis of the proposed method on ft10 instance

Method	Makespan
LSTM alone	1982
SS-LSTM without the self-supervised mechanism	1816
SS-LSTM	1613

5. Conclusion

This paper has proposed a novel and effective method named self-supervised LSTM for solving complex JSSP (see **Fig. 3**), which treated JSSP as a multiple subclassification problem. In

the proposed method, the detailed-level channel is for extracting detailed production states features while the system-level channel is for system-level feature extraction. Using those two channels with self-supervised LSTM can accurately deal with complex JSSP (see **Fig. 7**). Moreover, to make a trade-off between classification and feature extraction, the level of self-supervision is set to 0.4 (see **Table 2**). Each component's impact on the proposed method has been conducted from an ablation study (see **Table 3**). Significantly, the system-level channel reduced the make-span of 8.38%, and the supervised mechanism improved the performance of 11.18%, respectively. In summary, the proposed SS-LSTM could effectively and accurately solve complex JSSP.

References

[1] Y. Fu, J. Ding, H. Wang, and J. Wang, "Two-objective stochastic flow-shop scheduling with deteriorating and learning effect in Industry 4.0-based manufacturing system," Appl. Soft Comput. J., vol. 68, pp. 847–855, 2018.

[2] E. M. Frazzon, A. Albrecht, M. Pires, E. Israel, M. Kück, and M. Freitag, "Hybrid approach for the integrated scheduling of production and transport processes along supply chains," Int. J. Prod. Res., vol. 56, no. 5, pp. 2019–2035, 2018.

[3] K. Wang, H. Luo, F. Liu, and X. Yue, "Permutation Flow Shop Scheduling with Batch Delivery to Multiple Customers in Supply Chains," Ieee Trans. Syst. Man, Cybern. Syst., vol. 48, no. 10, pp. 1–12, 2018.

[4] H. W. Ge, L. Sun, Y. C. Liang, and F. Qian, "An effective PSO and AIS-based hybrid intelligent algorithm for job-shop scheduling," IEEE Trans. Syst. Man, Cybern. Part systems Humans, vol. 38, no. 2, pp. 358–368, 2008.

[5] L. Asadzadeh, "A local search genetic algorithm for the job shop scheduling problem with intelligent agents," Comput. Ind. Eng., vol. 85, pp. 376–383, 2015.

[6] G. R. Weckman, C. V. Ganduri, and D. A. Koonce, "A neural network job-shop scheduler," J. Intell. Manuf., vol. 19, no. 2, pp. 191–201, 2008.

[7] Z. Zang et al., "Hybrid Deep Neural Network Scheduler for Job-Shop Problem Based on Convolution Two-Dimensional Transformation," Comput. Intell. Neurosci., vol. 2019, no. ii, 2019.

[8] X. Shao, C. Soo Kim, and D. Geun Kim, "Accurate Multi-Scale Feature Fusion CNN for Time Series Classification in Smart Factory," Comput. Mater. Contin., vol. 65, no. 1, pp. 543–561, 2020.

[9] X. Shao and C. Kim, "Multi-step Short-term Power Consumption Forecasting Using Multi-Channel LSTM With Time Location Considering Customer Behavior," IEEE Access, p. doi:10.1109/ACCESS.2020.3007163, 2020.

[10] H. Fisher and G. L. Thompson, "Probabilistic learning combinations of local job-shop scheduling rules," Ind. Sched. vol. 3, no. 2, pp. 225–251, 1963., vol. 3, no. 2, pp. 225–251, 1963.

[11] S. Lawrence, "An Experimental Investigation of heuristic Scheduling Techniques," Suppl. to Resour. Constrained Proj. Sched., 1984.

Enhancing Environmental Data Forecasting Performance by Utilizing Multi-region Data with Hard-parameter sharing

Jang-Ho Choi, Miyoung Jang, Taewhi Lee, Jongho Won and Jiyong Kim
ETRI (Electronics and Telecommunications Research Institute)
Daejeon, South Korea
[e-mail: {janghochoi, myjang, taewhi, jhwon, kjy}@etri.re.kr]
*Corresponding author: Jang-Ho Choi

Abstract

Although deep neural network models are capable of learning complex non-linear relationship between input and target data, they require a large amount of well-balanced data in order to reach high performance level. Unfortunately, such abundant situations are quite rare in practice that in environmental data forecasting, for instance, datasets are not only severely imbalanced, but also scarce. Hence, this paper presents a multi-headed deep-neural network model that can effectively learn multi-region datasets mitigating data imbalance and insufficiency. The proposed architecture learns common features from multiple regions in addition to region-specific features of the target. The experimental studies show that the proposed network improves prediction performance by utilizing additional multi-region data more effectively.

***Keywords*:** Environmental Data Forecasting, Data Imbalance, Few-shot Learning, Water Quality Management

1. Introduction

In environmental engineering, simulation-based forecasting methods have been popular choices as they can accurately forecast impacts of multiple use case scenarios. However, such methods require valid assumptions, which must be designed by domain experts. Moreover, they often require extensive runtime computation to forecast. With remarkable research progress in deep learning, deep learning models have become appealing alternatives in various domains such as weather forecast, water quality management. Google Research, for instance, recently proposed machine learning models for precipitation forecasting that generate 0-6 hour forecasts in 1km resolution within 10 minutes [1].

Although deep learning models are capable of learning complex non-linear relationship between input and target data, they often require a large amount of data to discover optimal representations. However, datasets are not only scarce and expensive, but also often severely imbalanced in practice. Insufficient and imbalance datasets can lead to biased models, resulting poor prediction performance.

Furthermore, it is often the case that the users are more interested in rare events. In water quality management, for instance, algal bloom is the

This work was supported by the Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No. 2018-0-00219, Space-time complex artificial intelligence blue-green algae prediction technology based on direct-readable water quality complex sensor and hyperspectral image).

"interesting" event as it can be quite harmful to many living organisms, directly affecting the water quality. However, such events occur only once or a few times in several years regionally, making very difficult to learn repeating patterns. Combining regional data may increase the data volume in the minor classes, but unfortunately distant regions show different patterns due to diverse region-specific characteristics. In fact, we could observe performance degradation if models were trained with additional nearby datasets in raw.

Hence, in this paper, we propose a multi-headed deep neural network model that can effectively learn region-specific features in addition to common features. The proposed approach allows to mitigate data imbalance and insufficiency by utilizing additional data from other regions.

2. Modelling Environmental Data

In this section, we demonstrate characteristics of environmental data and model architecture to mitigate data imbalance and insufficiency.

2.1 Characteristics of Environmental Data

Environmental data, i.e. air quality and water quality data, are generally multivariate, time-series data. These data have become more abundant as network-equipped sensors and Internet of Things (IoT) are now widely available, which can collect and provide data and services rather automatically and autonomously [2]. However, even in the case of automatic collection, not only the cost of collecting data is quite expensive, but also, warning or catastrophic-level of events are very rare, occurring only a few times a year or less.

Another characteristic of environmental data is that they are generally collected from multi-sites. Even if these data consist of the same features, the underlying region-specific characteristics may vary. Hence, it is often not a good idea to utilize multi-region data in raw for local forecast.

2.2 Handling Data Imbalance

One of the popular data-level techniques is to generate new instances in training dataset. For example, Synthetic Minority Oversampling Technique (SMOTE) [3] is a simple and fast method to generate new instances of minorty classes using Euclidean distance vectors among nearest neighbors. Despite its simplicity, SMOTE works surprisingly well in practice [4]. Furthermore, deep-learning based approaches have recently become popular, generating new data points using variational autoencoder or generative adversarial networks.

2.3 Utilizing Multi-region Data

In environmental data forecasting, data augmentation by adopting nearby datasets seem to be a reasonable solution. However, we discovered that it is often not the case, as each regional dataset has its own unique regional characteristics. The comparison between training regional dataset and additional nearby dataset is shown in **Table 1**.

Table 1. Chlorophyll-a Prediction Results (JangGye) - Training Regional Dataset (JangGye) vs. Additional Dataset (JangGye, DaeCheong)

	R^2	MSE	Micro Acc.	Macro Acc.
JangGye	**0.5297**	**36.16**	**73.98%**	**66.72%**
+DaeCheong	0.4207	44.55	68.65%	49.04%
Difference	26.19%	-18.83%	7.76%	36.05%

In order to preserve region-specific characteristic, we adopt hard-parameter sharing from multi-task learning [5]. Hard-parameter sharing is a technique to place common hidden layers across multiple tasks, while keeping task-specific layers. Hard-parameter sharing is known to reduce overfitting, but it is generally restricted to tasks that are closely related one another [6].

Hence, we propose a multi-headed deep neural network model to extract general features from multiple regions, while learning region-specific features in separate layers. The proposed architecture is illustrated in **Fig. 1**.

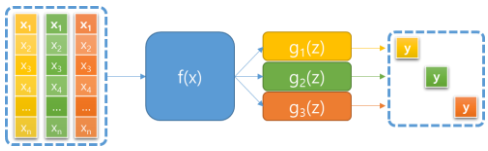


Fig. 1. Multi-headed Neural Network Model for multi-region data

3. Experimental Studies

In this section, we introduce experimental studies on the multi-headed neural network model for multi-region environmental datasets.

3.1 Datasets and Preprocessing

For the experimental studies, we used daily water quality data of four regions, namely JangGye, DaeCheong, CheongAm and NamGang. The datasets are collected by Korea's National Institute of Environmental Research (NIER) using water quality monitoring stations from 2012 to 2018. The water quality data consist of eight measurements: water temperature, pH, electrical conductivity, dissolved oxygen, total organic carbons, total nitrogen, total phosphorus and chlorophyll-a. The data imbalance is severe as shown in **Table 2**.

Table 2. Distribution of Water Quality Datasets (2012-2018) - Chlorophyll-A (Chl-A)

	Class 0	Class 1	Class 2	Class 3
Chl-A	0~15mg/m^3	15~25mg/m^3	25~100mg/m^3	100mg/m^3+
JangGye	1217	255	142	0
DaeCheong	1451	164	70	2
CheongAm	794	523	535	10
NamGang	1478	156	365	5

3.2 Experiment Result

In this study, we formulated problem to forecast the concentration of chlorophyll-a in three days. The primary goal is to investigate effectiveness of the multi-headed neural network in multi-region data setting, or any data that are in similar setting. Hence, we compared models in the same network setting: a simple two LSTM layer model. In the case of multi-headed model, the first layer is shared among the all four regions, while while separate second layer is assigned to each region.

Table 3. Performance Comparison (R^2, MSE, Micro and Macro Accuracy)

	R^2	MSE	Micro Acc.	Macro Acc.
Baseline (Single Region)	0.5297	36.16	73.98%	66.72%
Baseline (Multiple Regions)	0.4207	44.55	68.65%	49.04%
Multi-headed (Multiple Regions)	**0.5633**	**32.70**	**75.36%**	**70.04%**

The experiment result shows that the multi-headed model outperforms in all four metrics, R^2, MSE, Micro Accuracy and Macro Accuracy, as shown in **Table 3**.

4. Conclusions

In this paper, we present a multi-headed deep-neural network model, which can utilize multi-region datasets effectively, mitigating data imbalance and insufficiency. We showed that training additional region datasets in raw can degrade forecasting performance even when they are geographically close. Then, we demonstrated that the proposed model could utilize additional region datasets more effectively, showing improvements in several evaluation metrics. From the experimental studies, we could assume the proposed network sustains region-specific characteristics, while generalizes common parameters relatively well.

References

[1] Agrawal, S., Barrington, L., Bromberg, C., Burge, J., Gazen, C., and Hickey, J., "Machine Learning for Precipitation Nowcasting from Radar Images," arXiv:cs.CV/1912.12132, 2019

[2] Choi, J., Park, J., Park, H. and Min, O. (2017), DART: Fast and Efficient Distributed Stream Processing Framework for Internet of Things. ETRI Journal, 39: 202-212. doi:10.4218/etrij.17.2816.0109

[3] N. V. Chawla et al., SMOTE: Synthetic Minority Over-sampling Technique, Journal of Artificial Intelligence Research, Vol. 16, 2002.

[4] J. Choi, J. Kim, J. Won and O. Min, "Modelling Chlorophyll-a Concentration using Deep Neural Networks considering Extreme Data Imbalance and Skewness," 2019 21st International Conference on Advanced Communication Technology (ICACT) doi: 10.23919/ICACT.2019.8702027.

[5] Caruana, R. "Multitask learning: A knowledge-based source of inductive bias." Proceedings of the Tenth International Conference on Machine Learning. 1993.

[6] Sebastian Ruder, "An Overview of Multi-Task Learning in Deep Neural Networks," arXiv:cs.LG/1706.05098, 2017.

A Study on the Frame on News of VR and AR

Sang Hee Kweon and Hyeon-Ju Cha
Department of Social Science, Sungkyunkwan University
Seoul, South Korea
[e-mail: skweon@skku.edu, top1@skku.edu]
*Corresponding author: Hyeonju Cha

Abstract

This study empirically analyzed the evolution of the news coverage of the development of two technologies, Virtual Reality and Augmented Reality, by applying Frame Theory. The 'network' analysis was utilized that examines word choices, relationships between words, and overall logical structures created by the relationships in order to analyze the existence of news frames for each developmental stage (period) of Virtual Reality (VR) and Augmented Reality (AR) technologies. To this end, content analysis was conducted on VR and AR related 7,079 news pieces for the duration of five years (2012 to 2016). As a result, news about VR and AR displayed similar frames. The news about the two technologies placed their focus on the technical aspects from 2012 to 2014, while the focus was placed on the spheres of cultural life such as contents from 2015 to 2016. As such, the news during the initial developmental stage of VR and AR gradually started from technology-oriented reporting methods, then moved to the reporting methods focusing on cultural aspects related to the two technologies and their users.

***Keywords*:** Virtual Reality, Augmented Reality, Frame, Semantic Network.

1. Introduction

In recent years, along with Virtual Reality, Augmented Reality is also attracting attention. Virtual Reality (VR) allows to access the virtual world created by a computer program and to experience the senses such as vision and hearing as in the real world, while Augmented Reality (AR) presents a virtual object in the form of hologram to overlap an existing physical space. AR can be found in location-based services for smart phones, for instance, that add virtual information on real environments. The reverse is also possible – AR can add information in reality to a virtual reality environment. The <Pokémon Go> that surprised the world in July 2016 has emerged as a new game by adding AR to existing characters and stories.

Considering the features of such technologies, first, VR means a reality that is accepted as real even though it is not. VR supports its users to experience a distinct virtual reality that is separated from the actual reality. On the other hand, AR is a technology that shows the reality overwritten with virtual information. As <Pokémon Go>, AR interacts with GPS in the real-world or overwrites a Pokemon on camera. AR is known to be invented as an alternative to and overcome the technical limitations of VR; however, currently, it is as popular as VR.

The timings of initial development of the two technologies – VR and AR – differ. VR is not a completely new technology. Studies with attempts to implement a virtual reality had been continuously conducted since 1960s. However, the prerequisite conditions for implementing a virtual reality did not exist. Then a new

opportunity loomed as technologies such as computing power, communication speed, 3D sensing, and high-resolution display develop and with the declining cost of hardware devices. AR, on the other hand, was first used by Tom Caudell of Boeing in 1990. He tried to increase the effectiveness of work by describing the aircraft wiring assembly process via superimposing the virtual image on the actual screen, which allowed workers to easily distinguish between numerous wires connected to the headset monitor. Since the mid-2000s, it began to incorporate AR technology in civil services. Sony Computer Entertainment in 2007 had launched the AR game <The Eye of Judgment>, and in 2009, it scored a great success with <Eye Love Pet>. As such, VR and AR are being used in various parts of society that we are living. For instance, they are utilized in various fields including entertainment (games, movies, theme parks), service businesses (advertising, retail/shopping, tourism, health), education, media, industrial (automotive, manufacturing, real estate, and construction) [1] .

2. Theoretical Background

2.1 Study on the News Frame (Frame Theory)

The frame analysis in the present study is being conducted based on key words, subjects, the relationship between VR and AR technologies with social culture, and the direction of reporting that appear in news reporting process regarding VR and AR per period, in addition to Frame Theory. Gitlin [2] , a U.S. media researcher, argued that the mass media's reporting is captured by a 'frame' and such a 'frame' itself has ideological effects, using the concept of frame. Gitlin [2] defined 'frame' as "persistent patterns of cognition, interpretation, and presentation, of selection, emphasis, and exclusion." This means that the producers of the news they create specific topics and meanings according to the form of ownership, text, audience, and the form of the overall culture of a society. Thus, in the process of reporting the news on VR and AR, the press has a tendency of reporting with a specific view and frame. Further, the news on AR and VR and their structure can be divided into such two aspects. First, a thematic frame reports the national strategies, political events, and political issues related to VR and AR in general and abstract circumstances, rather than in terms of historical background or social structures. An episodic frame, on the other hand, focuses on specific cases or events in using AR or VR. In other words, the two are tendencies that appear in the developmental process of AR and VR technologies.

2.2 Research on Semantic Network Analysis and Centrality

Semantic network analysis is an application of social network analysis that emphasizes the relationship between actors in identifying social system structures on communication messages. While social network analysis focuses on the relationship that links each actor as the driving force, not on each actor, semantic network analysis that applies this method analyzes the meaning through the structural relationships between words, the components of the message, instead of the actors [3].

Semantic network analysis is a computer-based message analysis method that compensates the traditional content analysis method's shortcomings of not being able to rule out subjective interpretation of a researcher and its labor-intensiveness [4, 5]. In addition, there is an advantage that it can capture the dynamic meaning in the context of interacting relationships rather than the fixed property or concept of individual words [6]. Furthermore, semantic network analysis can extract the hidden structural patterns in a message by focusing on where each word is located within the overall network or how the words identified in the semantic network are creating what type of structure, moving away from being based merely on the combination between each word.

3. Research methods

3.1 Analysis Subjects and Data Collection

This study has collected data from media reports on VR and AR. Data collection was conducted through a news search on the nation's largest portal site, Naver. The extraction targets were comprehensive national daily newspapers. The selection of the newspapers included both conservative and liberal newspapers in order to

maintain the proper balance. Moreover, the analysis period of this study is from January 1, 2012 to December 31, 2016. From the media reports for five years, articles related to VR and AR were collected in addition to other data related to the two keywords. from 2012 to 2016, the amount of collected articles on VR is a total of 3,668 pieces, while the amount of data is 76,326 pieces. For AR, the collected amount for the five-year period is 3,411 pieces with the total amount of data as 40,158 pieces. Furthermore, [Figure 1] is a graph that illustrates the difference between Naver news data amount on VR and AR from 2012 to 2016. As depicted in [Figure 1], the data amount of VR was lower than AR in 2012 and 2013, while the data amount on VR had rapidly increased since 2014, which resulted in the amount of 25,507 pieces, twice the amount of data on AR in 2016.

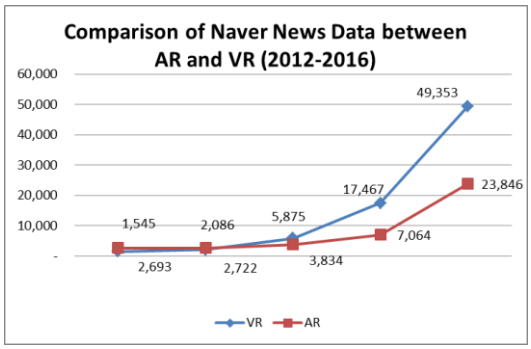


Fig. 1. 2012-2016 Trends in Naver news on AR and VR

3.2 Methods of Analysis

From the collected data, the process of Textom text cleaning was conducted by elimination of meaningless propositions, word endings, and punctuation marks. Frequency matrix analysis was also conducted. The process provides the matrix data composed of occurrence frequency among the purified words. In addition, the matrix data may proceed to further analysis by using the network analysis programs such as NodeXL.

All of the processes were processed on Textom, the big data solution.

The extracted words were then analyzed using the network analysis program NodeXL that visualizes the data on the relationship between key words and other words. NodeXL is a social network analysis tool designed to collect, analyze, and visualize the different types of network data. NodeXL can collect data through a built-in importer from various online social media data such as Facebook, Twitter, Youtube, Flickr, and Wikipedia. Also, NodeXL can calculate key figures of social network analysis, including degree, centrality, density, and communitying coefficient, while it can express the networks subject to analysis in various visualization diagrams through its own internal visualization function.

To illustrate the details of the analysis method, the keyword, 'virtual reality,' collected based on the period of 2012 will be used. First, the collected data after collection, cleaning, and frequency matrix on Textom were loaded to NodeXL to progress with Graph Metrics calculation. Graph Metrics can describe the entire network or describe the subgroups or characteristics of a particular destination within one network. Next, in the auto fill columns in NodeXL's visual attributes menu, the values of edges and vertices were set to conduct mapping on visual characteristics. First of all, from edges columns, the entire edge weight was selected from the edge color, width, opacity, and visibility; specifically, the edge color was set at 128,128,128 in the color RGU 192,192,192. Edge width was set at 1.0 to 5.0, and the edge opacity and edge visibility were aimed at understanding and identifying the relationship between key words via the descending sort order. Therefore, the smallest value of the edge opacity and edge visibility was set to 50. Additionally, in the vertices columns and vertex color, size, visibility, the entire vertex degree centrality was selected. All options relating to these were set to their default values.

3.3 Research questions

Specific research questions are as follows:

Research Question 1: What differences in importantly treated words in relation to VR and AR exist for each period?

Research Question 2: How the key words on VR and AR in the media structure semantic networks in each period?

4. Conclusions

Table 1. Analysis of semantic networks an centrality of keywords on VR per year

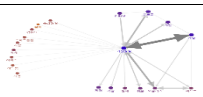
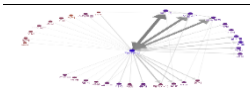
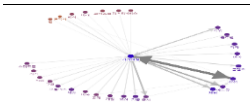
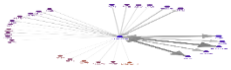
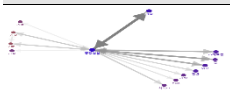
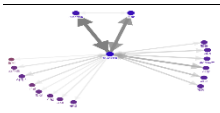
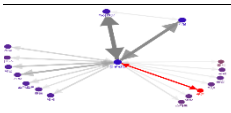
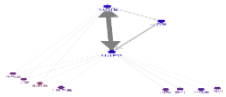
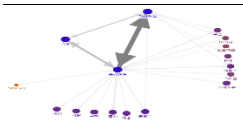
Keywords on VR Networks by Year	Description of Network Graph
	The semantic networks related to VR in 2012 keywords shows 'VR' at the center. Also, the terms such as 'game', 'global', 'technology', and 'development' are linked to VR by thick lines, meaning that they constitute the center of the overall network. With Graph Metrics calculation, specific subjects or subgroups were subject to analysis that created one big community around VR in 2012. VR has numerous words as its child. Still, such words are mostly related to technology.
	In 2013, two communities of centers appeared compared to 2012. First, the community centering around VR is the largest community that linked to 'game', 'global', 'Oculus', and 'development' with think lines. Another community is centered around 'game', with 'development', 'global', 'lift', and 'Internet' at its lower ranks. Also, it is notable that VR started to appear amongst the orange colored words in 2013 semantic networks graphs.
	The semantic networks of keywords related to VR in 2014 show 'VR', 'headset', 'Oculus', and 'Samsung Electronics' at the center. Among the communities, the community centered around VR is the largest, while a small number of words are connected to the community centered around the other three words, 'headset', 'Oculus', and 'Samsung Electronics'. Also, 'headset', 'Oculus', and 'Samsung Electronics' show a high level of connectivity with VR.
	In 2015, the keyword 'Samsung' has a high level of connectivity with VR, followed by 'headset', 'experience', 'game', and 'release'. Two communities can be found in the keywords semantic networks related to VR in 2015. The first is surely the largest community around VR, followed by the community of 'Samsung'. Yet, the words connected with the keywords of the community around 'Samsung' are not many with a low level of connectivity strength.
	The last is the semantic networks of keywords related to VR in 2016. The network consists of one community. At the center is VR, with 'AR', 'technology', 'experience', 'contents', 'Korea', 'game', and 'global' are located at the lower level. Among those words, 'AR', 'technology', and 'experience' have the thickest connecting lines

Table 2. Analysis of semantic networks an centrality of keywords on AR per year

AR related Keywords Networks by Year	Description of Network Graphs
	On the left is the semantic networks of keywords related to AR in 2012. In this graph, the keyword AR is located at the center. Also, the word 'technology' is connected with the centering word AR with a very thick line. Other words such as 'smartphone', 'app', and 'use' are also connected to the centering word AR with a line in the second thickness.
	The patterns appear in 2013 semantic networks have AR at the center with two other words, 'VR' and 'technology', connected with thick lines showing a high level of relevance. Also, 'development', 'experience', and 'game' had a high level of connectivity. Although AR is the center of one large community, another small community has been created such as 'VR' and 'technology'.
	Looking at the semantic networks related to the keyword AR in 2014, the word AR is at the center of the network. Additionally, the words like 'VR', 'technology', 'development', 'Google', and 'experience' are connected with AR with thick lines, showing that they form the center of the overall networks. With the Graph Metrics calculation, small communities or certain subjects were analyzed. Accordingly, in 2012, one large community was formed around AR, and another small community was formed around the words like 'VR' or 'technology'. AR also had numerous words under it.
	The keywords with a high level of connectivity to the keyword AR are 'VR' and also 'technology' in 2015. In 2015, the keywords semantic network on AR shows two communities. The first one is surely the largest community around AR, followed by one centering around the keyword 'VR'. Still, the words connected with the keywords of the community around 'VR' are not many with a low level of connectivity strength.
	Looking at the semantic networks related to the keyword AR in 2016, the words 'AR' and 'VR' are at the center of the network. With the Graph Metrics calculation, small communities or certain subjects were analyzed. Accordingly, in 2016, one large community was formed around AR. And as in 2015, another community was formed around the word 'VR'. Yet, such a community has a more complex pattern compared to the year before, and the word 'I-Navi' had an independent community.

The first research question in this study is aimed at examining the differences in words with significance in the media in their coverage of VR and AR per period. First, the highly ranked words mentioned by Naver news per year, the frequency of the word, "virtual reality," increased from 1164 to 1798 times from 2012 to 2016. Looking at each year in detail, the high ranked key words of 2012 were 'VR (1164),' 'game (233)', 'technology (145)','global (130)', 'development (108)' and 'internet (105)'. In 2013, 'VR (1177)', 'game (329)', 'oculus (206)', 'global (167)', 'internet (141)' and 'development (132)' – mainly technical terms occupied the top of the ranking. The word that was mentioned most frequently in 2014 was also 'VR (1676)', followed by 'oculus (261)', the user tool for VR contents. In the third place, 'headset (251)' was ranked a term that means the same. Additionally, 'Samsung electronics(210)' showed the frequency of over 200 times. In 2015, the frequency of 'VR' was 1881 times, followed by 'Samsung' with 318 times. Next, the words such as 'experience (193)', 'headset (176)', 'game (172)', 'launch (139)' and 'contents (127)' appeared. Finally, in 2016, 'VR' with the frequency of 1798 times was ranked as the number one, followed by 'experience (222)', 'AR (222)', 'technology (186)'.The second research question in the present study is identifying how the keywords related to VR and AR and their semantic networks differ in the media reports by period. Specifically, by using the keywords, it analyzes how VR and AR form public knowledge. The second research question is to analyze the semantic networks of words based on the top keywords analysis in the research question 1. Results of the second research question are described in [Table 1] below.

References

[1] Hwang-Su Jeon, Mi-Kyung Han, Jong-hyun Chang, "The Journal of Trends Analysis of the Electronics and Telecommunications," *Electronics and Telecommunications Research Institute*, Vol 32, No. 1, pp. 93-101, 2017.

[2] Gitlin. T. Berkeley, "*University California Press*," 1980.

[3] Wasserman, Stanley, Faust, Katherine. "*Structural analysis in the social science*," New York : Cambridge University Press, xxxi, 825 p. : ill. ; 23 cm, pp. 756-801, 1994.

[4] Han-Woo Park, Loet Leydesdorff, "Korean Data Analysis Society," *Journal of the Korean Data Analysis Society*, Vol 6, pp. 1377-1387, 2004.

[5] Ha-Yong Jung, "Korean Journal of Communication & Information," *Communication & Information Society*, Vol 16, No. 2001, pp. 108-135, 2001.

[6] Sang-bae Kim, "The Korean Journal of International Relations," *Korean Association of International Studied*, Vol 51, No. 4, pp. 51-77, 2011.

A Unique Approach towards Database Service Recovery using Centrally Managed Backup Containers

Adnan Younas and Hussein A Mubaireek
Computer Operation Department, Saudi Aramco
Dhahran, Saudi Arabia
[e-mail: adnan.younas@aramco.com, hussein.mubaireek@aramco.com]
*Corresponding author: Adnan Younas

Abstract

This paper discusses the data readiness which is critical to the overall application availability, and as a whole supports and maintains the critical business scenario. It is crucial to design and implement strategies for high data availability and reliability using the unique backup recovery procedures. Backup and recovery are one of the most critical and frequently performed operations especially during the forensic analysis when the data is ceased for any modification. The intent is to create a corrective backup architecture to mitigate both human and technical weaknesses. The idea is to enhance the database transactions backup and data recovery processes with well-defined frequency that facilitate continuous data protection. Also, support wider recovery scope and data flashback capabilities using Network Attached Storage (NAS).

***Keywords*:** Data Recovery, High Availability, Disaster Recovery, Forensic Analysis

1. Introduction

Currently, a complete standby database system needs to be kept up and running in a standby mode to apply changes. This entails maintaining a separate database installation for standby database. The maintenance procedures include frequent patching, monitoring, and administration hosted in high-end hosting platform.

- In case of a disaster where the production database is lost, standby database can be activated and used instead. In the meantime, a copy from standby database is required to rebuild production environment.

- In case of a disaster where both production and standby databases are both lost, a backup restore from tape media is required in order to recover the database [1]. This is usually a very lengthy process since backup tape are relatively slow.

- Once the standby database is activated and opened for read/write, it cannot be switched back to standby mode again. It will need to be completely rebuilt again [2]. Reconstruction of Volumetric Models

2.1 Proposed Model

There are two part of the proposed model one is done at the storage level and the second part is handled at the database management system level. The Production database backups are being kept in a shared NAS or SAN storage area. A copy from offline redo logs for 7 days are also kept in the same storage area as well. There is no need to have a dedicated running standby database instance [3]. The proposed model involves the two tier approach, at tier one

addresses the problem at the storage level whereas the tier two focuses on the overall architecture which includes database, storage and backup restoration side of it.

At tier 1 level, focus would be on storage i.e., creation of new volume (which requires a small amount of space for metadata) from a snapshot of a volume. The system instantly replicates data volumes and datasets as transparent, virtual copies-true clones-without compromising performance or demanding additional storage. This process allows you to get older version from one of your databases hosted on the same volume without affecting other databases. Benefits at this level comprised of cost reduction, storage footprint, and complexity of environments that support new-product development, software and penetration testing, or deployment of virtual server infrastructures.

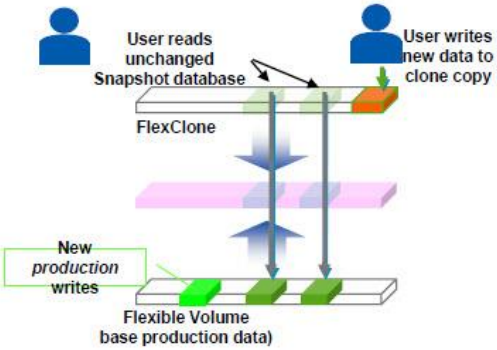


Fig. 1. NetApp FlexClone (adopted:BCC Services [3])

Tier 2 mainly discuss the architecture and functionalities on top of Tier 1 level. This represents the overall model functionalities.

In case of a disaster where the production database is lost, a new host can be provisioned that have the database binary pre-installed and ready. Then the database can be started using "boot from backup" mode using the backup files stored in the shared storage area. The database can be started up to any timeframe required [4].

Another option: the production database can be immediately recovered up by booting the database instance from backup files in the shared storage. In case of a disaster where both production and standby databases are lost, a new host can be provisioned and started up in very short time using the backup files from storage area. There is no need to restore database backup from tape.

In case of logical data corruption, (e.g. end user deletes a record from a table), the object can be retrieved from the backup files to the production database.

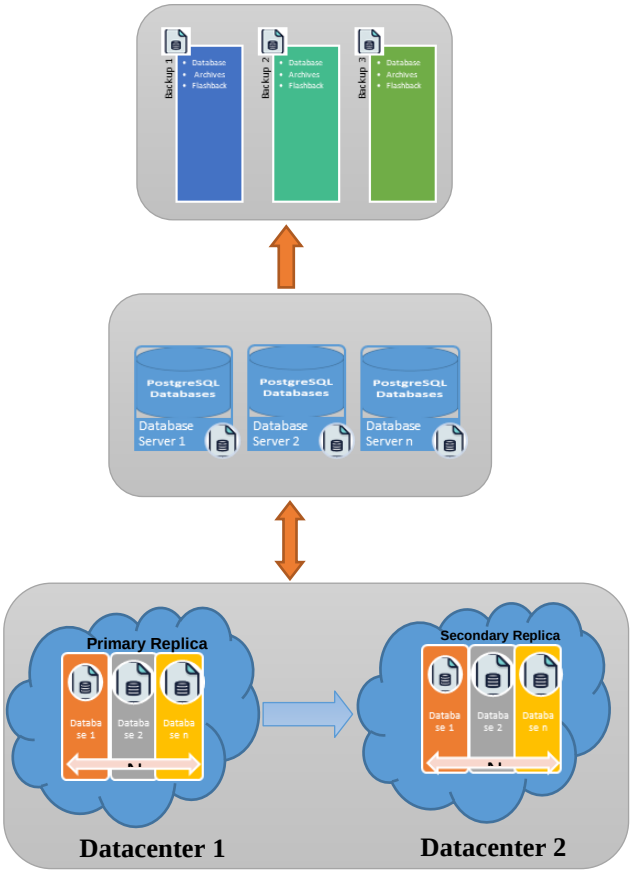


Fig. 2. Proposed Data Recovery Model

2.2 Model Dynamics on different scenarios

During the course of this model implementation different scenarios have been considered. For example, user accidentally or maliciously deletes or modifies several records from one or more key tables on a production database.

An application bug that changes data in a database to incorrect or inappropriate values.

Other scenarios could be malware that modifies or wipes data, encrypted the data. It reduces the cost such as hardware (Server, CPU, Memory and SAN Storage), resource management such as patching, vulnerability assessment and upgrades

etc. It also helps during the forensic events when data is ceased for any modification.

2.3 Impact and cost of the model implementations

During the course of model implementation, storage used approximately 2-3 Terabytes NAS storage hosted on primary datacenter Cloud for primary replica. Similar amount of storage is required for secondary replica on secondary datacenter cloud. The backed-up transactional data are compressed and the size estimation is an average based, it depends upon database activity and protection retention requirement.

3. Model Dynamic on different scenarios

Different dynamics have been discussed in this paper in relation to the model such as, user accidentally or maliciously deletes or modifies several records from one or more key tables on a production database. Another scenario could be an application bug that changes data in a database to incorrect or inappropriate values. It also addresses the possibility of malware that modifies, wipes data or encrypt the data. Another aspect considered was the cost, such as hardware (Server, CPU, Memory and SAN Storage), Resource management such as patching, vulnerability assessment and upgrades etc. It would also be useful during forensics event when data is ceased for any modification [5].

4. Implementation

PostgreSQL database management system has been used during the implementation and testing of the above-mentioned model. Create Online Full Database Backup and a container (folder/directory) on NAS storage

For example: /nas_pg_replica

Then mount the NAS storage to the operational PostgreSQL Database instance

Schedule a job to perform an online full database backup, For example:

```
pg_basebackup -h localhost -p $PGPORT -U $USER -Fp -Xs -P -v -R -D /nas_pg_replica/$PG_NAME -T /pg/data/ tablespace_name=/nas_pg_replica/$PG_NAME /tablespace_name
```

Sync the Database Activity Changes (sync-wal)

Schedule a loop to copy local generated write-ahead logs to the NAS based container. This could be scheduled based on the database activity for example every 10 minutes. For stricter RPO, this could be customized to lesser value.

```
/bin/rsync -avz /pg/local_archive/replica/ /nas_pg_replica/PG_NAME/wal
```

Recover & Startup PostgreSQL Replica (from NAS storage backup

Pre-requirement: from a host that has the PostgreSQL software packages

Update the PGDATA environmental variable to point to the NAS storage container
For example: export PGDATA=/nas_pg_replica/PG_NAME

Option #1: Startup the Replica without roll-forward (as snapshot of the base replica backup)

Startup the PostgreSQL database instance
```
pg_ctl -s -D /nas_pg_replica/PG_NAME start -w
```

Option #2: Startup the Replica with roll-forward (use the wal files to recover the database)

Modify the $PGDATA/recovery.conf to
```
restore_command = 'test ! -f %f && cp /nas_pg_replica/PG_NAME/wal/%f "%p"'
recovery_target_time = '07/26/2020 23:30:00.0 +03'
```

Data and time can be adjusted as needed on the above code.

Startup the PostgreSQL database instance in recovery mode.
```
pg_ctl -s -D /nas_pg_replica/PG_NAME start -w
```

Once the recovery is completed the recovery file $PGDATA/recovery.conf is changed to $PGDATA/recovery.done

5. Conclusions

This paper has discussed and presented an interactive method to manipulate the proposed model based on different scenarios such as sudden data loss, data corruption or even in case of physical disaster such as loss of datacenter or communication between datacenters. Protection and reversal of data tampering incident has also been considered and tested. It has been noticed that there was a significant reduction of Recovery Point Objective (RPO) to the least due to the proposed model. On the other hand, it also drastically improves the Recovery Time Objective (RTO). This is due to the database transactions exist on disk, which eliminate the tape restoration overhead. Proposed model also helps in avoiding computing cost of marinating one-to-one hot standby replica (server provisioning, CPU, Memory, SAN storage, and associated licenses) Avoid operational cost of marinating one-to-one hot standby replica (patching, upgrades, vulnerability scans, syncing, downtime, changes) Expose multiple cost-effective failover scenarios.

It is possible to keep the timeframe between viable backups within a few minutes. The only cost associated with this is the cost of the storage required to keep the backups. Facilitate flashback data for one or two weeks

References

[1] Anastasia Lubennikova, "Advanced PostgreSQL backup & recovery methods," *pgconf.eu*, 2018.

[2] PostgreSQL Installation and Maintenance, Swiss Federal Institute of Technology, D-BSSE Wiki, 2015.

[3] http://www.bccservices.com/infrastructure-solutions/netapp/netapp-platform-os/netapp-flexclone/, accessed on 16th July 2020.

[4] EnterpriseDB® Corporation, "EDB Postgres™ Backup and Recovery Guide", 2019.

[5] https://www.enterprisedb.com/blog/maintaining-postgresql-for-high-performance-what-is-wrong-or-right-what-consider, accessed on 10th August, 2020.

Comparative Evaluation of State-of-the-Art Approaches in Association Rules Mining

Syed Mazhar Ali Shah and Dr. Anazida Zainal
School of Computing, Faculty of Engineering, University Technology Malaysia (UTM)
Skudai, Johor Bahru, 81310, Malaysia
[syed.mazhar@graduate.utm.my, anazida@utm.my]

Abstract

Association Rules Mining (ARM) is one of the most popular but expensive activity in pattern discovery. Numerous ARM approaches (techniques, models and frameworks) have been introduced in the past few decades, but most of these have been implemented using a minimum support threshold. This is because of either data dependency between database transactions or a convenient way forward to deal with the 'curse of dimentionality'- the exponential growth of search space- on medium or large-scale systems, with enough efficiency within existing resources. Recent reviews are mainly focused on ARM approaches that need setting of minimum threshold values to reduce search space, without including other 'supportless' approaches that do not need such minimum threshold values. Also, only positive correlation between transactions are considered while analyzing these recent ARM approaches. The basic purpose of this study is to introduce an algorithmic efficacy oriented categorical perspective of recent ARM approaches that not only include the conventional ARM approaches but also the missing logic-based supportless ARM approach. Algorithmic Efficacy factors like rare rules mining, negative rules mining, and dealing with database scarcity with relevant challenges are briefly identified. As for the result, it seems that ARM approaches for pattern discovery are still broadly open for further improvements along with inclusion of new mining approaches, with subsequent comparison for an improved categorization and comprehensiveness, with respect to algorithmic efficacy.

Keywords: Automatic Association Rules Mining, Negative Association Mining, Supportless Rule Mining, Supportless Negative Association, ARM, Minimum Support Threshold, Logic based Rule Mining

1. Introduction

Association Rule Mining (ARM) has revolutionized nearly every application domain for its ability of expressing associations among items or events occurring together, thus bringing forth new knowledge for enlightened decision making. The prime challenge to this problem area is the 'curse of dimensionality', where growth rate of associations is exponential with increasing number of items or events. Therefore, different pruning strategies were introduced to reduce the search space while performing association rule mining on frequently occurring items or events (frequent itemset mining-FIM), mostly considering their positive correlations, only. The same was applied to mine the infrequent ones (exceptions/rare), also with positive correlations taken into consideration. Frequent as well as infrequent itemsets were explored for positive correlations, in the early evolutionary phases of association rule mining, on small binary and multi-valued transactional databases, and is called frequent rule mining [1] and exception/rare rule mining [2] respectively. Such frequent as well as rare rule mining, that only considers the positive correlations and apply minimum thresholds like 'support' and 'confidence' as primary pruning strategy to control exponential growth of search space, is generally known as 'positive association rule mining'. Likewise, when negative correlations

like absence of an event with absence or presence of another event are considered too, and implemented in the realm of ARM then such ARM approaches are negative ARM approaches [3][4]. How a particular ARM approach deals with the problem of missing values in a considerable portion of the database, provides another algorithmic efficacy factor (*or Algorithmic capability parameter*) for ARM approach categorization.

2. Association Rule Mining (ARM)

After the introduction of positive association rule mining in [1] many algorithms of this family were developed. All such algorithms have two common steps: a) construction of a set of all candidate itemsets and, b) extraction of frequent association rules from each candidate itemset and application of an interestingness measure to each extracted frequent association rule. The mining of the set of all the candidate itemsets, is a recursive process of building supersets from previously constructed subsets or vice versa, with a single or multiple database-scans. These two steps are also valid in mining rare itemsets [2].

3. State of-the art ARM approaches

The main purpose of this study is to emphasis the supportless ARM approach in the over-shadowing traditional support-based ARM approaches, to fully ascertain the state of the art in ARM [29]. It has been observed that the main limitation of the early supportbased ARM approaches is not only limited to the loss of information due to high minimum support values but also these approaches fail to consider the indirect proportionality- negative correlation- in the relationships between attributes or items within the database. Negative supportless ARM approach is a solution to overcome these limitations associated with positive only ARM approaches. Based on this, the major ARM approaches have been categorized as under:

3.1 Positive Association Rules Mining (PARM) Approaches

In the perspective of transactional databases, each attribute/item can either be present or abset. No other state is possible like in categorical databases. In such binary state database scenario, if we consider the relationship of a presence of an item with the presence of another item only, without any consideration of any possible relationship of the absence of one, with the absence or presence of other, then, we are only considering the positive associations. The following techniques are representative of this approach:

3.1.1 Generate-and-Test (Apriori Approach)

Introduced by [1] Apriori approach is well known for its wider application in the early development of ARM techniques. It identifies the frequently occuring unique items in the database and combines these unique items to form larger itemsets till they cease to appear frequently in the database. This traditional approach is a mature field of research in Positive Association Rules Mininng that uses a generate-and-test strategy to produce a candidate set of frequent itemsets starting from a set of 1-length frequent itemsets to be subsequently used to produce n-length frequent itemsets where $n>1$: $n \varepsilon N$ till no more set of $n+1$-length frequent itemsets is produced further.

3.1.2 Divide and Conquer (FP tree approach)

Another approach in Positive Assocaition Rules Mining is based on divide-and-conquer approach. In this approach, the database is divided into smaller chunks depending on the need at hand [5]. Also, the database is compressed and mapped into a tree structure to infer frequently occuring intemsets. The FP-Growth algorithm [6] is the representative algorithm of this approach that provides a replacement of devide and conquer Apriori approach used to identify the most frequently occuring itemsets by maintaining a data stack of tree datastructure to control search space. The FP-Growth algorithm is far more efficient than a conventional Apriori algorithm variant due to its tree implemenation for itemset search [7].

3.1.3 Lattice based Approaches

Lattice-Based Association (LBA) rules mining approaches try to deal with the problem of repetitive generation of candidate set of itemsets

in their traditional predicessors. LBA rules mining approaches are composed of two phases: 1. Frequent Itemset Lattice building 2. Generation of association rules from the lattice generated. Lattice based approaches are much more efficent as compared to the traditional approaches due to elimination of repetitive candidate itemsets generation through multiple database scans. Once a lattice is generated then the same can be used with different minimum thresholds with no need of re-running the algorithm for different minimum thresholds [8].

3.2 Negative Association Rules Mining (NARM) Approaches

An ARM approach that considers negative correlations between the events represented by the database along with the positive ones to mine the association rules, is categorized as a NARM approach. It is important to note at the outset that all approaches that are categrorized as a NARM approach, is also a PARM approach but the reverse is not true. All NARM approaches not only produce negative associatoin rules but also produce their positive counterparts too. But a PARM approach once categorized so, only produces the positive association rules. Many variations of the NARM approach are as follows:

3.2.1 Collective Strength Based Approaches (CSBAs)

Originally proposed by [9], this approach and its variants focus on the actual attribute correlation among items rather than absolute statistical presence in the database to evaluate and find large itemsets. CSBAs are efficient, statistically robust and can handle large databases. Relative measure used in such ARM approaches is suitable for negative ARM in both evenly or unevenly dense databases.

3.2.2 Heuristic Based Approaches (HBAs)

Heuristics are a problem-solving approach to use shortcuts for non-optimal but satisfactory solutions, given a limited time and resources. Thus, heuristics are a flexibility technique for quick decisions, particularly when working with large and complex data. For example, while indexing we can use heuristics of similarity for grouping similar transactions according to a similarity criterion to minimize the number of comparisons in ARM. Many performance metrices have been checked against heuristic approach in NARM to ascertian the pros and cons of a particular heuristic based forecasting approaches [10].

3.1.3 Logic Based Approaches (LBAs)

Any logical reasoning is used to establish a criteria to prune the search space with or without a need for minimum support [11]. The main idea in logic-based approach is to take advantage of relational query languages and data mining primitives to represent data mining tasks as object-relational queries [30]. Logic-based database languages help to represent, maintain and utilize high-level knowledge in LBAs.

4. Comparative Evaluation

The evaluation criterion has been recognized and depicted then the outcomes and results are displayed in the wake of applying the evaluation criteria into existing approaches. As the aim of this study is to provide a categorization and comparison of the most prominent positive and negative state-of-the-art approaches in determining how far each ARM approach supports the discovery of association rules from a transactional database efficiently. Thus, the gained insight is shared as potential algorithmic efficacy factors to be considered for these ARM approaches.

4.1 Evaluation Results

This section describes and clarifies concisely the evaluation criteria used in this study. To indicate the existence and absence of the desired efficacy factors in the discussed state-of-the-art ARM approaches, we have performed analysis against each approach and has been formulated in tabular form as presented with signs "✓" and "x".

Table 1.1 Efficacy Factors Evaluation results of state of-the-art ARM approaches

Efficacy Factors / Approaches	Mines RARE Rules?	Mines –VE Rules?	Supportless ARM?	Sparsity Handling?
Lattice Based	✓[18]	x	x	✓◆[27]
Collective Strength	✓[19]	✓[9] [17]	x	x
Heuristic	✓[16]	✓[12]	x	x

Forecasting				
Logic based	✓[20]	✓[21]	✓[14]	✓◆[28]
Apriori	✓[22]	✓[4]	×	✓[25]
FP Growth	✓[23]	✓[24]	×	✓◆[26]
◆	Needs enhancement to deal with sparsity			

As shown in **Table 1.1**, there are many limitations and weaknesses in the mentioned ARM approaches with respect to the algorithmic efficacy factors named 'Rare Rule Mining', '-Ve Rules Mining', 'Supportless Rules Mining', and 'Database Sparsity Handling'. Although many researchers have identified different algorithmic efficacy factors of different ARM approaches but no study has presented a comparison of these ARM approaches with respect to these algorithmic efficacy factors as proposed in this paper to the best of our knowledge.

4.2 Potential Efficacy Factors for Future Association Rules Mining Approaches

The comparative evaluation of the considered ARM approaches led us to the insight of potential areas where the algorithmic capability in such approaches are improved by focusing on new efficacy factors. A few such potential efficacy factors are discussed as follows:

4.2.1 Associatoin Rules Mining Approach Improvement for Better Completeness of Negative Rare Rules Set

'Completeness of the mined candidate rare rules set' of an algorithm, is the ratio of mined unique negative rare rules to those mined by an exhaustive search algorithm, from the candidate set of itemsets [12]. Thus, completeness provides a quantification of how much of the domain knowledge latent within the Database is being captured in the form of negative rare itemsets/rules. Greater the ability of an algorithm to mine every exceptionality within the Database thus, is a potential algorithmic efficacy factor worthy being considered while opting a particular ARM approach. It is important that this potential algorithmic efficacy factor is dependent on 'execution efficiency', and 'memory resources' of an ARM approach. Therefore, a persistent tradeoff between this potential algorithmic efficacy factor and 'execution efficiency' can be observed in many algorithms [13]. Despite this interdependency, if an ARM approach facilitates better 'completeness of the mined candidate rare rules set' even on the expense of decreased 'execution efficiency' within eccepted limits, is a considerable ARM appraoch improvement.

Furthermore, the supportless approach in negative rare rule mining through 'Implication Equivalence' in propositional logic [14][15], favors the selection of the most important rules set as 'coherent rules' but majority of the rules failing the strict criteria of 'equivalence' are ignored. This strict criteria indicates a potential room for further improvement in the ARM approach with respect to completeness.

Therefore, another supportless ARM approach seems eminent with a better logic to increase the degree of completeness of the mined candidate rare rules set.

4.2.2 Association Rules Mining Approach Improvement for Supportless Identification of Minimum Prunning Condition

All traditional approaches including the lattice-based ones, we need a suitable minimum threshold value (i.e. minimum support/confidence) to facilitate early control of exponential expansion of search space [1][8].
But, application of minimum thresholds (as a pruning strategy) needs setting appropriate minimum threshold values for 'support' and 'confidence'. Finding such appropriate minimum threshold values has always been challenging, as these minimum threshold values are subjective and may not be accurate for a certain application domain scenario [14]. Therefore, other 'supportless' approaches must be considered that are free of such minimum threshold values in the algorithmic design. The ARM approach where such a minimum threshold solely depends upon the used database, reduces chances of information loss. Similarly, dealing with sparsity 'supportlessly' is an additional research challenge with respect to ARM approach, and inclusion of such supportless approaches in the overall analysis along with the traditional support-based ARM approaches, could only present the absolute state of the art.

4.2.3 Association Rules Mining Approach Adoptation for Enhanced 'Exceptionality Measure'

An exceptionality measure is a formula (function), designed to quantify the 'exceptionality' (interestingness with respect to an application domain) of a 'supportless negative rare itemset', using information both within and out of the database to be mined. Usually, the 'out of the database' information is composed of input from 'domain expert' in the form of setting some starting variable values [16]. For an efficient supportless ARM apoproach, exceptionality measure is imporotant because it ensures the quality of the mined negative exception rules set. The exceptionality measure used to quantify the 'relevance of the candidate itemset to the application domain' leaves profound effects on the accuracy of the 'supportlessly' mined 'negative rare rules set'. Furthermore, not only the final interesting itemsets are decided by the exceptionality measure but exceptionality measures are also used to reduce the search space, specially, when negative rare itemsets are being pruned [17].

Thus, exceptionality measure not only enhances the accuracy of the 'supportlessly' mined negative rare rules set but also plays an important role in increasing the execution efficiency of a particular algorithmic design.Therefore, a small improvement in the design of the exceptionality measure can greatly increase the accuracy of the information represented by the final 'negative rare rules set'.That is why, a future ARM approach needs 'fine-tuning' of the well-established exceptionality measures by using all the possible information available, preferably 'within' the database.

5. Conclusions

This study aims at providing an overview of state of-the-art ARM approaches ensuring that the supportless approach is included to add new perspective. Also, many of such approaches have been compared for being comprehensive with respect to preselected evaluation criteria in order to categorize these as being negative or positive only ARM approaches. Although, these positive and negative association approaches are not constrained to prove required outcomes, but are analyzed against certain efficacy factors as performance criteria. We conclude that there is a need for further research in the field of ARM approaches, especially in the areas mentioned as potential efficacy factors. Also, for the future work, we aim at developing a new ARM approach having the indicated potential efficacy factors.

References

[1] Agrawal, R., T. Imieliński, and A. Swami. *Mining association rules between sets of items in large databases*. in *Proceedings of the 1993 ACM SIGMOD international conference on Management of data*. 1993.

[2] Zhou, L. and S. Yau. *Association rule and quantitative association rule mining among infrequent items*. in *Proceedings of the 8th international workshop on Multimedia data mining:(associated with the ACM SIGKDD 2007)*. 2007.

[3] Antonie, M.L. and Zaïane, O.R., 2004, June. An associative classifier based on positive and negative rules. In Proceedings of the 9th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery (pp. 64-69).

[4] Ramasubbareddy, B., Govardhan, A. and Ramamohanreddy, A., 2011. Classification based on positive and negative association rules. *International Journal of Data Engineering,(IJDE)*, *2*(2), p.84

[5] Wang, X.-Z., R.A.R. Ashfaq, and A.-M. Fu, *Fuzziness based sample categorization for classifier performance improvement.* Journal of Intelligent & Fuzzy Systems, 2015. 29(3): p. 1185-1196.

[6] Han, J., J. Pei, and Y. Yin, *Mining frequent patterns without candidate generation.* ACM sigmod record, 2000. 29(2): p. 1-12.

[7] Siahaan, A.P.U., A. Ikhwan, and S. Aryza, *A Novelty of Data Mining for Promoting Education based on FP-Growth Algorithm.* 2018.

[8] Le, T. and Vo, B., 2016. The lattice- based approaches for mining association rules: a review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *6*(4), pp.140-151.

[9] Aggarwal, C.C. and P.S. Yu, *Mining associations with the collective strength approach.* IEEE Transactions on Knowledge and Data Engineering, 2001. 13(6): p. 863-873.

[10] Ghafari, S.M. and C. Tjortjis, *A survey on association rules mining using heuristics.* Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2019. 9(4): p. e1307.

[11] Kumbhar, S., et al., *Pattern Discovery Using Apriori and Ch-Search Algorithm.* International Journal of Computational Engineering Research, 2015. 5(3).

[12] Ghafari, S.M. and Tjortjis, C., 2019. A survey on association rules mining using heuristics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(4), p.e1307.

[13] Sadhasivam, K.S. and T. Angamuthu, *Mining rare itemset with automated support thresholds.* Journal of Computer Science, 2011. 7(3): p. 394.

[14] Sim, A.T.H., et al., *Logic-based pattern discovery.* IEEE Transactions on Knowledge and Data Engineering, 2010. 22(6): p. 798-811.

[15] Sim, A.T.H., M. Indrawan, and B. Srinivasan, *The importance of negative associations and the discovery of association rule pairs.* International Journal of Business Intelligence and Data Mining, 2008. 3(2): p. 158-176.

[16] Chowdhury, K.P., 2019. Supervised Machine Learning and Heuristic Algorithms for Outlier Detection in Irregular Spatiotemporal Datasets. *J. Environ. Inform*, *33*.

[17] Kadir, A.S.A., Bakar, A.A. and Hamdan, A.R., 2014. Frequent positive and negative (FPN) itemset approach for outlier detection. *Intellignt Data Analysis*, *18*(6), pp.1049-1065.

[18] Zhang, J., Jiang, Y., Chang, K.H., Zhang, S., Cai, J. and Hu, L., 2009. A concept lattice based outlier mining method in low-dimensional subspaces. Pattern Recognition Letters, 30(15), pp.1434-1439.

[19] Mukhi, S.N., 2010. A confidence-based aberration interpretation framework for outbreak conciliation. *Online Journal of Public Health Informatics*, *2*(1).

[20] Hamamoto, A.H., Carvalho, L.F., Sampaio, L.D.H., Abrão, T. and Proença Jr, M.L., 2018. Network anomaly detection system using genetic algorithm and fuzzy logic. *Expert Systems with Applications*, *92*, pp.390-402.

[21] Duggirala, R. and Narayana, P., 2013. Mining Positive and Negative Association Rules Using CoherentApproach. *arXiv preprint arXiv:1308.2310*.

[22] Hanguang, L. and Yu, N., 2012. Intrusion detection technology research based on apriori algorithm. *Physics Procedia*, *24*, pp.1615-1620.

[23] Aung, K.M.M. and Oo, N.N., 2015, March. Association rule pattern mining approaches network anomaly detection. In *Proceedings of 2015 International Conference on Future Computational Technologies (ICFCT'2015) Singapore* (pp. 164-170).

[24] Wang, B., Chen, D., Shi, B., Zhang, J., Duan, Y., Chen, J. and Hu, R., 2017. Comprehensive association rules mining of health examination data with an extended FP-growth method. *Mobile Networks and Applications*, *22*(2), pp.267-274.

[25] Ye, F.Y., Wang, J.D. and Shao, B.L., 2005, August. New algorithm for mining frequent itemsets in sparse database. In *2005 International Conference on Machine Learning and Cybernetics* (Vol. 3, pp. 1554-1558). IEEE.

[26] Tan, J., Bu, Y. and Yang, B., 2009, August. An efficient frequent pattern mining algorithm. In *2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery* (Vol. 2, pp. 47-50). IEEE.

[27] Lin, T.Y., Hu, X.T. and Louie, E., 2003, March. Using attribute value lattice to find closed frequent itemsets. In *Data Mining and Knowledge Discovery: Theory, Tools, and Technology V* (Vol. 5098, pp. 28-36). International Society for Optics and Photonics.

[28] Vishnu Priya, R. and Vadivel, A., 2012. Partition-based sorted pre-fix tree construction using global list to mine maximal patterns with incremental and interactive mining. *International Journal of Knowledge Engineering and Data Mining*, *2*(2-3), pp.137-159.

[29] Patel, B.M., Bhemwala, V.H. and Patel, A.R., 2018. Analytical Study of Association Rule Mining Methods in Data Mining.

[30] Liu, H.C., Zeleznikow, J. and Jamil, H.M., 2006, January. Logic-based association rule mining in XML documents. In *Asia-Pacific Web Conference* (pp. 97-106). Springer, Berlin, Heidelberg.

RuleLat: A Rule-based Model Visualization Tool for Machine Learning Interpretability

Josue Obregon[1] and Jae-Yoon Jung[1,2*]
[1]Department of Industrial and Management Systems Engineering, Kyung Hee University, South Korea
[2]Department of Big Data Analytics, Kyung Hee University, South Korea
[e-mail: jobregon@khu.ac.kr, jyjung@khu.ac.kr]
*Corresponding author: Jae-Yoon Jung

Abstract

Rule-based models used for classification are typically simple and interpretable compared to more complex models such as support vector machines, ensemble models and neural networks. However, the classification ruleset may become difficult to interpret when the number of rules as well as the number of conditions on each rule grow large. In this paper, we present a simple way to visually represent a rule-based classifier called RuleLat. The generated model visually presents the relationships among predicted classes, rules and antecedents in a modified lattice that is simpler to analyze and understand. Additionally, an output example is presented to point out the benefits and applicability of the lattice of rules.

***Keywords*:** decision rules, formal concept analysis, classification, interpretable machine learning

1. Introduction

The interpretability of machine learning models in many scenarios is as important as its prediction accuracy [1]. One simple way to ensure the interpretability of a classification models is to use methods that are intrisincally more interpretable by the human intuition. Generating rule-based classifiers are a popular option because interpreting them is straightforward, that is, a set of if-then rules that gives a final prediction if the conditions are satisfied [2].

However, the interepretability of a rule-based models can be affected if the number of rules or the number of conditions inside a rule is too large. Moreover, if conditions are repeated many times, it becomes more difficult to keep track of the decision mechanisms of the model and its interepretability decreases.

In this short paper, we present a visual tool, called RuleLat, which summarizes and presents in a rule lattice the relationship among rules, conditions and predicted classes. The construction of the lattice is based on the formal concept analysis technique [3].

The remainder of this paper is structured as follows, Section 2 gives the neccesary background to construct the rule lattice. Section 3 presents an output example to validate the method, and Seciont 4 conclude the paper.

2. Conversion of Ruleset to Lattice

This work was supported by the Korea Evaluation Institute of Industrial Technology (KEIT) grant funded by the Korean government (MOTIE) (No. 20008690) and the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2017H1D8A2031138).

2.1 Rule-based Classifiers

A rule-based classifier is a set of rules, generally expressed in the form of an implication r: $\boldsymbol{A} \rightarrow y$, where $\boldsymbol{A}$ is a logical conjunction of conditions $a \in \boldsymbol{A}^*$, and $\boldsymbol{A}^*$ is the set of all possible conditions derived from the training set $\boldsymbol{D}=\{(\boldsymbol{x}_i, y_i)\}$ with i = 1 to m, where m is the number of data. $\boldsymbol{A}$ is also called a rule antecedent or the left-hand side (LHS) of the rule, and y is the consequent class label, also called the right-hand side (RHS) of the rule.

It is said that a rule $\boldsymbol{r}$ covers a record $\boldsymbol{x}_i$ if all the conditions in $\boldsymbol{A}$ satisfy the attributes of $\boldsymbol{x}_i$. Decision rules can be grouped into rulesets $\boldsymbol{R}$ to collectively classify all the training examples. There are two types of rulesets: unordered and ordered. Unordered rulesets evaluate all rules and the predictions of the triggered rules are combined with a defined mechanism such as voting or averaging. Ordered rules, also known as decision lists, evaluate the rules in order, and the prediction of the first triggered rule is taken as the predicted class label [4]. In this paper, we work with decision lists.

To measure the performance of a ruleset $\boldsymbol{R}$, each record of the training sett $\boldsymbol{D}$ is evaluated to check if it triggers at least one rule. In this way, the coverage and accuracy of a rule over a dataset D can be calculated as:

$$cov(r) = |\boldsymbol{A}| \,/\, |\boldsymbol{D}|$$
$$acc(r) = |\boldsymbol{A} \cap y| \,/\, |\boldsymbol{A}|$$

where $|\boldsymbol{A}|$ is the number of records that satisfy the LHS of r, $|\boldsymbol{D}|$ is the total number of records in the training set, and $|\boldsymbol{A} \cap y|$ is the number of records that satisfy both the LHS and the RHS of r.

2.2 Formal Concept Analysis

The formal context analysis (FCA) is a mathematical framework which helps to formalize the notion of concept and conceptual hierarchy [3]. FCA takes a collection of objects and attributes so then it derives formal contexts and formal concepts, making possible to construct concept lattices to make it easier to visualize the hierarchy of concepts. Because of those reasons, it is ideal to depict the relations among rules and conditions.

Formally, a context $\boldsymbol{K}=(\boldsymbol{G},\boldsymbol{M},\boldsymbol{I})$ is conformed with two sets $\boldsymbol{G}$ and $\boldsymbol{M}$, and a relation I between them. The elements $g \in \boldsymbol{G}$ are called objects, and the elements $m \in \boldsymbol{M}$ are called attributes of the context. The relation between an object g and an attribute m is denoted $g\ \boldsymbol{I}\ m$ or $(g,m) \in \boldsymbol{I}$ and it is read as "the object g has the attribute m" [3].

A context can be represented in a cross table in which its columns represent attributes, and its rows represent objects. A cross in the column g and row m of the table represents the relation $(g,m) \in \boldsymbol{I}$. An example can be seen in Figure 1a. For a set $\boldsymbol{E} \subseteq \boldsymbol{G}$ the set of all attributes common to the objects in $\boldsymbol{E}$ is defined as:

$$\boldsymbol{E}^{\uparrow} = \{m \in \boldsymbol{M} \mid g\ \boldsymbol{I}\ m \text{ for all } g \in \boldsymbol{E}\}$$

In the same way, for a set $\boldsymbol{F} \subseteq \boldsymbol{M}$ the set of all objects having in common the attributes in $\boldsymbol{F}$ is defined as:

$$\boldsymbol{F}^{\downarrow} = \{g \in \boldsymbol{G} \mid g\ \boldsymbol{I}\ m \text{ for all } m \in \boldsymbol{E}\}$$

A formal concept of the context $\boldsymbol{K}$, is a pair $(\boldsymbol{E},\boldsymbol{F})$ with:

$$\boldsymbol{E} \subseteq \boldsymbol{G},\ \boldsymbol{F} \subseteq \boldsymbol{M},\ \boldsymbol{E}^{\uparrow} = \boldsymbol{F} \text{ and } \boldsymbol{F}^{\downarrow} = \boldsymbol{E}$$

The set $\boldsymbol{E}$ is called the extent and the set $\boldsymbol{F}$ is called the intent of the concept $(\boldsymbol{E},\boldsymbol{F})$. Moreover, having two concepts $(\boldsymbol{E}_1,\boldsymbol{F}_1)$ and $(\boldsymbol{E}_2,\boldsymbol{F}_2)$, the formal concepts are naturally ordered by the sub-concept and super-concept relation when:

$$(\boldsymbol{E}_1,\boldsymbol{F}_1) \leq (\boldsymbol{E}_2,\boldsymbol{F}_2) \Leftrightarrow \boldsymbol{E}_1 \subseteq \boldsymbol{E}_2\ (\Leftrightarrow \boldsymbol{F}_2 \subseteq \boldsymbol{F}_1)$$

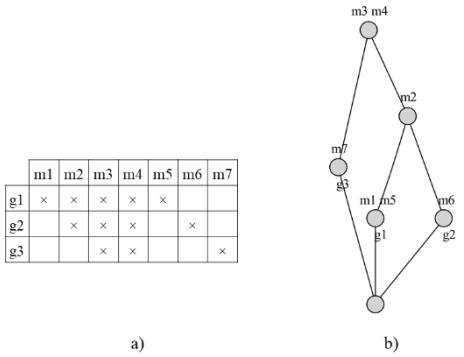


Fig. 1. (a) Cross table and (b) concept lattice example

Thus, a hierarchical order can be generated as a graphical representation called a concept lattice. An example of a concept lattice can be seen in **Fig. 1**. The previous definitions are enough to

understand how to generate RuleLat from a set of decision rules.

2.3 RuleLat

Using the FCA setting, a lattice of rules can be derived from a rule-based classifier.

A rule context $\boldsymbol{K_r} = (\boldsymbol{R}, \boldsymbol{A}, \boldsymbol{I})$ consists in the sets $\boldsymbol{R}$ and $\boldsymbol{A}$, and a relation $\boldsymbol{I}$ between them. The elements $r \in \boldsymbol{R}$ are the rules belonging to a ruleset, and the elements $a \in \boldsymbol{A}$ are called the antecedents of the rule context. The relation between a rule r and an attribute a is denoted $r\,\boldsymbol{I}\,a$ or $(r, a) \in \boldsymbol{I}$ and it is read as "the rule r has the antecedent a".

A rule context can be represented in a rule-antecedent cross table in which its columns represent antecedents, and its rows represent rules. Additionally, all the values of the class label y are added as columns in the rule-antecedent cross table, and each rule has a cross in the respective predicted class of the rule. The reason for this is because the antecedents and rules are grouped in one sub-concept for each class. A cross in the column r and row a of the table represents the relation $(r, a) \in \boldsymbol{I}$.

From the rule-antecedent cross table, a rule lattice can be derived. The rule lattice is a visual representation of the hierarchies among classes, antecedents and rules belonging to the rule-based classifier.

To add value to the rule lattice, it is annotated with meaningful information about the rules and the covered instances from the training set. Each edge is annotated with the number of instances covered by all the connecting lower neighbor nodes. Additionally, each rule node is annotated with its coverage and accuracy. The final lattice, called RuleLat, is a more interpretable visual representation of the relationship among rules, antecedents and predicted classes.

3. Example

For proving the validity of the proposed approach, a small example is presented in this section. The dataset used is a modified version of the new thyroid dataset with data of five lab tests used to predict a patient's thyroid class. In this dataset, class "1" denotes that a patient suffers from hyperthyroidism and class "-1" denotes that a patient is either normal or suffers from hypothyroidism.

The first step is to train a model, which was done using XGBoost (eXtreme gradient boosting) algorithm. For transforming and simplifying the result, the RuleCOSI method is used [5] which combine and simplify the result of an emsemble of decision trees into a single rule-based classifier. The simplified ruleset $\boldsymbol{R}$ can be seen in **Fig. 1(a)**.

a)

r_1: (Thyroxin <= 12.65) ^ (T3resin > 95.5) → -1
r_2: (Thyroxin <= 12.65) ^ (TSH_value > -0.15) → -1
r_3: (TSH_value <= 1.05) ^ (Thyroxin > 12.65) → 1
r_4: (TSH_value > 1.05) ^ (Thyroxin > 12.65) → -1
r_d: ( ) → 1

b)

	a_1	a_2	a_3	a_4	a_5	a_6	*1*	*-1*
r_1	x	x						x
r_2	x		x					x
r_3				x	x		x	
r_4					x	x		x
r_d							x	

c)

a_1: Thyroxin <= 12.65 a_2: T3resin > 95.5
a_3: TSH_value > -0.15 a_4: TSH_value <= 1.05
a_5: Thyroxin > 12.65 a_6: TSH_value > 1.05

d)

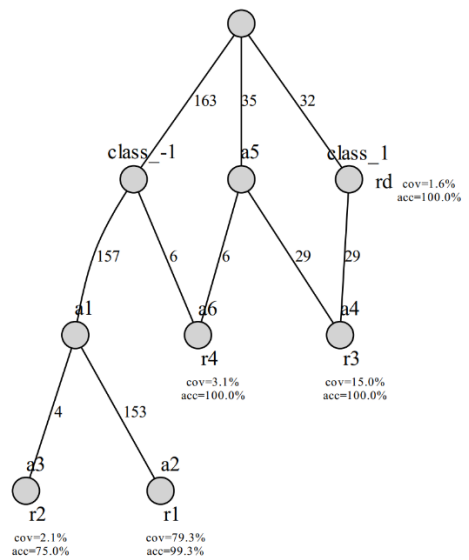


Fig. 2. Example of the RuleLat generated for the New Thyroid dataset. (a) Rule-based classifier $\boldsymbol{R}$, (b) rule-antecedent cross table, (c) map of antecedents, and (d) RuleLat lattice

Fig. 2(b) presents the rule-antecedent cross table which is constructed by relating the antecedents and classes with the 5 rules belonging to the simplified ruleset. The annotated rule lattice RuleLat is shown in Figure 2d along with its map of antecedents in **Fig. 2(c)**. RuleLat edges give information about the number of instances covered by the extent of the rule formal context on each node. On each rule node, the coverage and accuracy are displayed, which are calculated using the equations presented in Section 2.1.

Interpreting the diagram is simple and straightforward. The diagram is clearly divided in two subgroups of nodes that represent the two classes of this problem. Patients who suffer the hyperthyroidism are predicted with r_3 and the default rule r_d. The remainder of the cases are predicted with r_1, r_2 and r_4. Around 80% of the cases are covered by r_1 having 99.3% of accuracy. An interesting point of this example is condition a_5, which is used to predict both classes. This denotes that a_5 needs to interact with other conditions to be useful.

4. Conclusions

In this paper, we presented a method to generate visual representations of rule-based classifiers in the form of a rule lattice. The method is based on the formalisms of FCA. Additionally, one example is presented to show how the method works with real data.

As a future work, experiments with bigger rulesets can be performed and a more deep analysis of the capabilities offered by FCA to investigate the relationships among rules and conditions can be performed.

References

[1] Bibal, A. & Frénay, B. "Interpretability of machine learning models and representations: an introduction," In *Proc. of the 24th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, 2016.

[2] Tan, P.-N., Steinbach, M., & Kumar, V. *Introduction to Data Mining*, Boston: Pearson Addison Wesley, 2006.

[3] Ganter, B. & Wille, R. *Formal Concept Analysis: Mathematical Foundations*, Springer-Verlag Berlin Heidelberg, 1999.

[4] Fürnkranz, J., Gamberger, D., & Lavrač, N., *Foundations of Rule Learning*, Springer Science & Business Media, 2012.

[5] Obregon, J., Kim, A., & Jung, J.-Y.. "RuleCOSI: Combination and simplification of production rules from boosted decision trees for imbalanced classification." *Expert Systems with Applications*, pp. 126, pp. 64-82, 2019.

Comparative Study of Database Tools for Android Application: A Bird's Eye View

Hina Hussain[*], Dr. Nawab Muhammad Faseeh Qureshi[§], Dr. Qasim Ali[*], Dr. Isma Farah Siddiqui[*¥]
[*]Department of Software Engineering, Mehran University of Engineering & Technology Jamshoro, Pakistan
[§]Department of Computer Education, Sungkyunkwan University, Seoul, South Korea
[e-mail: engrhina15@outlook.com, isma.farah@faculty.muet.edu.pk[¥]]
[¥]Corresponding Author

Abstract

The database is the backbone of any software application. The prominent performance of the Database is a significant factor in the development and efficient execution of android applications. Some android apps need rapid input response, other processes much information, and without the support of an efficient database management system, this is nearly incomprehensible. While talking about mobile devices, then offline or local databases came into the picture. In this paper, the main focus is to compare the performance of available offline tools/libraries which are SQLite, Room Persistence Library, and ObjectBox in different scenarios. Firstly, perform CRUD operations on all three databases and calculate time in seconds and compare the time of each operation. Secondly, calculate the size of each database in MB. For this comparison, an android application has been developed to test the performance and efficiency of each tool. The evaluation results show that the Object Box library's performance is very well in terms of Create, Update, and Delete operations, and in Read operation, whereas the SQLite is better than the other two. In terms of size Room Persistence Library and SQLite, both are good than ObjectBox.

***Keywords*:** Android, database, SQLite, Room Persistence Library, Object Box, CRUD, Performance

1. Introduction

The smartphone industry is consistently developing and growing in terms of market size as well as in models and suppliers.

Right now, there are billions of Android cell phones. Android is an open-source Linux based operating system developed by Open Handset Alliance (OHA's), led by Google, and other companies. According to market research, Android has a 72.92% market share.[1] The number of accessible applications in the Google Play Store from December 2009 to October 2020 is 3.04 million.[2] As the market grows, the hardware is becoming more powerful, and the possibilities of data processing from cell phones are expanding. Most Android applications utilize some form of database to store particular user's data and other types of information.

A database can be considered as the backbone of any application. The performance of the database is significant in the construction of any application. Android offers built-in database functionality in the form of SQLite but there are a lot of libraries support database management. Some of them are Room Persistence Library, Object Box and Realm, and many more.

The main problem is to discover the current best solution on Android to store information on the device. Saving data in a database is an ideal solution for repeated and structured data for offline usage. However, most of the time, the developer community does not know which database library will be best according to their application requirements.

The main goal of paper deals with the analysis of the performance of these three database libraries in terms of most basic CRUD operations, as well as in terms of memory size. This comparison will help developers to choose the best database library to increase the performance of their apps. Section 2 focuses on the literature review and research to deal with databases and performance impacts. Section 3 gives a brief overview of database libraries. Section 4 describes the method and procedure. This section also introduces an application, called **DBToolInspector**, which performs testing and performance measurement and conducts results. Performance tests and their results are discussed in section 5.

2. Literature Review

Prior studies have been found that there are many research work is available for database comparison. A.Gupta, S.Tyagi, N.Panwar and S. Sachdeva [3] did a crucial study and compared different types of NoSQL Databases based on functional and non-functional features.Four major classifications of the NoSQL database are graph databases, document stores, key-value stores, and wide column stores. Then for additional investigation, one database is chosen from everyone and is contrasted with their data model, CAP theorem, distributive properties, and other factors. A.Talha KABAKUŞ [4] did a performance comparison of two widely used android databases, SQLite and Firebase. For this comparison, he develops an android application to execute common CRUD operation, and experimental results indicate that SQLite has better performance than Firebase except in delete operation.

S.Agarwal and K.Rajan [5] compare an SQL database (PostGIS) with a NoSQL database (MongoDB) for spatial and aggregate queries.

The tests they evaluate show that the NoSQL database gives better execution by a normal factor of 10x-25x which increments exponentially when the number of information increments in both classified and non-classified operations. N.Leavitt [6] discusses a brief history of SQL and NoSQL databases close by the contrasts between them. He noticed that SQL and NoSQL databases are complementary since they are intended to overcome various types of issues. N.Obradovic, A.Kelec and I. Dujlovic [7] did a performance analysis of the Android SQLite database. Their study's primary goal is to investigate the performance of the SQLite database in different scenarios, such as performing CRUD operations on encrypted or unencrypted data. Further, they discuss simultaneous access to the database.

3. Overview of Selected Database tools

3.1 SQLite

SQLite is an open-source SQL database that stores data to a text file on a device. Android comes in with built-in SQLite database implementation. SQLite is the most used database engine in the world. [8]

3.2 Room Persistence Library

Room is an ORM, Object Relational Mapping Library, which provides an abstraction layer over SQLite to allow easy database access. It is a part of Android Jetpack and created By Google Developers.

3.3 ObjectBox

ObjectBox is a super-fast database and synchronization solution, built uniquely for Mobile and IoT devices [9], and created by GreenRobot.

4. Methods and Procedure

Benchmarking of these databases tools performance is very complicated because it depends on different parameters ,such as device hardware ,OS version,file system and alse the dataset.Therefore this paper is not focus on external factors and datasets but focused on database operations.These test cases have been conducted on real device Oppo A5s whose specifications are listed in **Table 1**.

Table 1. Device Specifications

Android Version	API Level	RAM	CPU
Android 8.1.0 (Oreo) 8.	27	2GB	Octa-core (4x2.3 GHz Cortex-A53 & 4x1.8 GHz Cortex-A53)

An android application named **DBToolInspector** has developed to perform CRUD operations on all three databases. A Dataset used from google data sets which are "World Cities Population and Location" consists of 19 columns. DBToolInspector app extracts data from the CSV file and performs CRUD operations on that datasets. We have

used different sizes of data sets such as 15000, 30000, 50000, 100000, and test cases are executed 10x to calculate the average time in secs and size in MBS. The graph has been shown to analyze time in each CRUD operation and size in MB. **Fig. 1** shows the working flow of the DBToolInspector app. **Fig. 2** shows the interface of the android application.

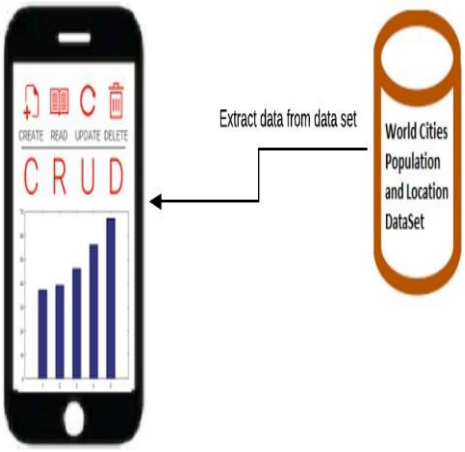


Fig. 1. Working flow of the app

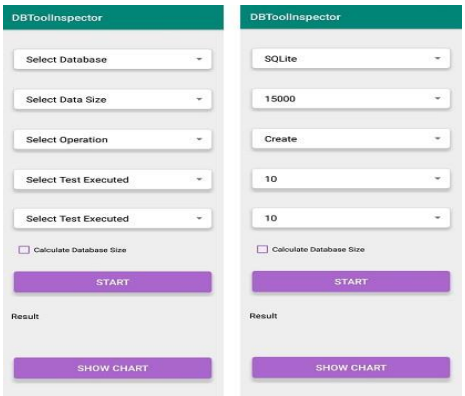


Fig. 2. DBToolInspectorApp interface

5. Results and Discussions

By performing the CRUD operations, we have achieved the following results.

5.1 Create

Fig. 3 shows that while saving data the best result is ObjectBox. Furthermore, the worst case is an SQLite database, and the Room also gives a reasonable result.

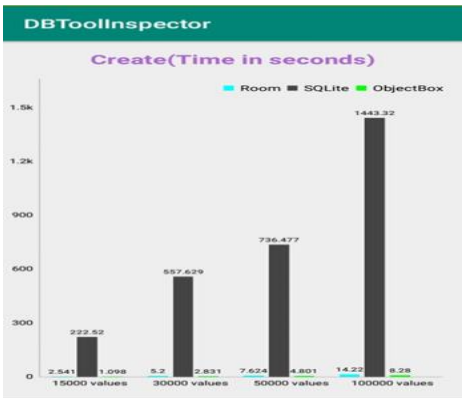


Fig. 3. Create Operation

5.2 Read

Fig. 4 shows that when we are reading data, the SQLite is incomparably faster than the other DB systems. Room and ObjectBox take more time than SQLite.

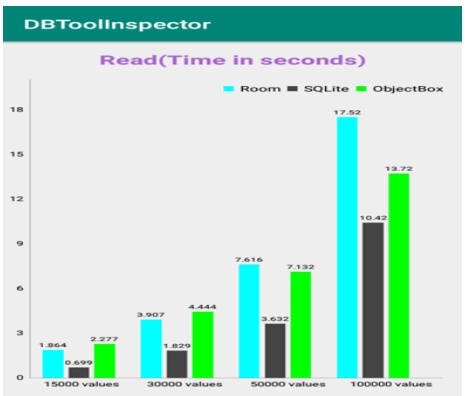


Fig. 4. Read Operation

5.3 Update

When updating database information, again, ObjectBox has beaten SQLite, and SQLite has the worst result, and Room Persistence Library has satisfactory results. The results can be observed in **Fig. 5**.

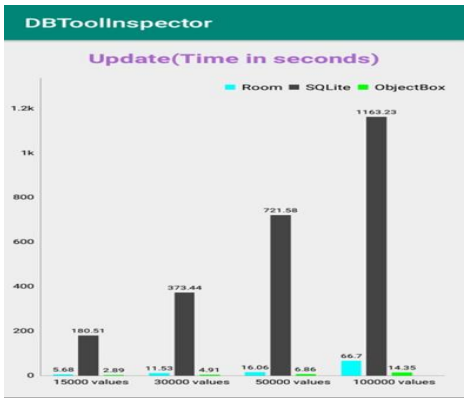


Fig. 5. Update Operation

5.4 Delete

For deleting data in Room, and SQLite has performed very similar results. Again ObjectBox is the best, but it should be noted that the database cleanup surgery is rarely used. The results can be observed in **Fig. 6**.

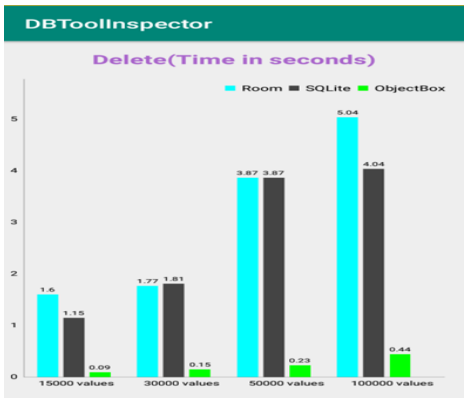


Fig. 6. Delete Operation

5.5 Size

Comparing the presented systems, attention should be paid to the memory used by the databases. In this case, the worst is Object Box, whose database size is 40.48 MB. Room Persistence Library and SQLite have achieved almost the same results as shown in **Fig. 7**.

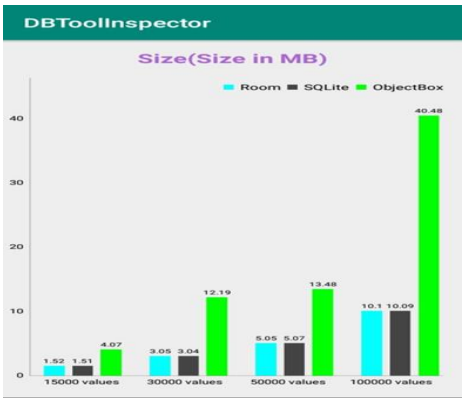


Fig. 7. Size

6. Conclusions

The main goal of this paper is a comparison of the database tool's performance in terms of CRUD and size. Evaluation results show that ObjectBox performance is very good while performing create, update, and delete operations, but in terms of size, it has the worst results. Moreover, when performing a read operation, SQLite performance is far better than the other two. Room Persistence Library performs well in all operations. In terms of size, both SQLite and Room Persistence Library have satisfactory results.

As the next step in research, our goal is to conduct more complex operations such as add more tables, perform joins and subqueries on that data, and then compare results. However, this does not mean that the result will be the same for complex queries.

References

[1] StatCounter Global Stats. N.d. Mobile Operating System Market Share Worldwide | Statcounter Global Stats. [Online] Available at :< https://gs.statcounter.com/os-market-share/mobile/worldwide>.

[2] Statista. N.d. Google Play Store: Number of Apps | Statista. [online] Available at:

[3] <https://www.statista.com/statistics/266210/number-of-available-applications-in-the-google-play- store/>.

[4] A.Gupta, S.Tyagi, N.Panwar and S. Sachdeva, "NoSQL Databases: Critical Analysis and Comparison," *Minjar Cloud Services Private Limited*, 2017.

[5] A.Talha KABAKUŞ "A Performance Comparison of SQLite and Firebase Databases from a Practical Perspective," *DÜZCE ÜNİVERSİTESİ, MÜHENDİSLİK FAKÜLTESİ*, 2019.

[6] S.Agarwal and K.Rajan "Analyzing the performance of NoSQL vs. SQL databases for Spatial and Aggregate queries*," Free and Open Source Software for Geospatial (FOSS4G)*, Vol.17, Article 4,2017.

[7] N.Leavitt, "Will NoSQL Databases Live Up to Their Promise? "*Computer*, 2010.

[8] N.Obradovic, A.Kelec and I. Dujlovic "Performance analysis on Android SQLite database," *18th International Symposium INFOTEH- JAHORINA*, 2019

[9] Sqlite.org. n.d. SQLite Home Page. [online] Available at:<https://www.sqlite.org/>.

[10] ObjectBox. 2020. Objectbox - Edge Database for Mobile and Iot. [online] Available at:<https://objectbox.io/>

Hyperledger Fabric for Livestock Traceability: Addressing the Impact of Poultry Supply Chain on Human Health and Climate Change

Aliyu Ibrahim, Cen Senfeng, Usmonov Kamoliddin, Ju Ho Kim, Akromjon Obidov and Chang Gyoon Lim*
Department of Computer Engineering, Chonnam National University
Yeosu, South Korea
[e-mail: {187282 ; 207906 ; 175120 ; 126494 ; 187126 ; cglim}@jnu.ac.kr]
*Corresponding author: Chang Gyoon Lim

Abstract

Poultry contributes significantly to the human food security and nutrition, protein, essential micro-nutrients and the provision of energy through the ability to convert a wide variety of agri-food wastes and byproducts into edible meat and eggs in short production cycles. Globally, the growth in demand for poultry meat and eggs is projected to be 121% and 65%, respectively, with Asia as the driving force. While the poultry sector has seen tremendous success and huge potentials ahead, public concerns about the sector's impact on the environment and human health, its role in climate change is triggering the need to embrace innovative technology to accurately record, track and trace information at various level of production in the sector so as to formulate accurate policies to address the issues. Thus, this paper proposes Hyperledger Fabric for the poultry management system. The system is proposed to manage the history of poultry rearing from a hatchery, farm, processing plant to distribution and retail outlet for customers. The objective is to manage the entire history of poultry products. The Hyperledger technology offers an accurate, decentralized, immutable and consensus process that promotes trust and transparency between stakeholders, and also facilitate easy identification of the sources of contamination in case of recall to safeguard public health.

***Keywords*:** Poultry Management System, Hyperledger, Blockchain, Poultry Traceability, Smart Contract

1. Introduction

Poultry contributes significantly to the human food security and nutrition, protein, essential micro-nutrients and the provision of energy through the ability to convert a wide variety of agri-food wastes and byproducts into edible meat and eggs in short production cycles [1]. It is one of the most dominant meat and food industries in the world with about 23 billion birds worldwide in 2016, i.e., approximately 3 birds per person in the universe [2]. Poultry meat and products are particularly the most dominant source of animal food in the livestock sector because it uses natural resources to supply protein acceptable by diverse cultures, traditions and religions around the world- estimated global production of eggs and meat is put to about 73 and 100 million tons, respectively [2]. Moreover, animal-sourced foods are essential for nutrition, particularly in children, pregnant women and the elderly, as they provide essential micronutrients – such as riboflavin, vitamin A and B-12, iron, calcium

and zinc- which plant sources are inadequate in providing the required quantities [1].

In the last 50 years, the demand for an animal source of food has been in the increased with poultry meat seeing an average growth of 5% annually [3]. The high demand for the poultry products is expected to continue with an increase in population growth, income and urbanization- the global population is expected to hit 9.6 billion by 2050, with approximately 70% of the population living in the urban area, while income is expected to rise by 2% per year [2]. Globally, the growth in demand for poultry meat and eggs is projected to be 121% and 65%, respectively, with Asia as the driving force. By 2025, The production of meat in East and South Asia is estimated to expand by 1.8 Mt with poultry meat as part of the bulk of the expansion [1]. In egg production, East Asia is projected to have the highest gains with 108% growth.

While the poultry sector has seen tremendous success and a huge potential ahead, public concerns about the sector's impact on the environment and human health, its role in climate change is triggering the need to embrace innovative technology to accurately record, track and trace information at various level of production in the sector so as to formulate accurate policies to address the issues. The various levels of production include poultry raising, processing, packaging and delivery. A study shows that food safety is the main concern to consumers, which comes before affordability and nutritional content [4]. Likewise the increasing trend in the lifestyle of consumers for frozen, ready-to-eat and package food are also shaping the demands on poultry product to be transparent [5-7]. To improve consumer confidence, trust and ensure food safety and easy recall (in case of disease outbreak), there is a need to ensure that recorded information from the source of livestock products through all the stages of rearing/production, processing and distribution is traceable and verifiable. Therefore, research and application of appropriate distributed ledger technology are essential to the poultry industry.

Blockchain technology which is underpinned by the Industry 4.0 is envisaged as the potential solution to alleviate the above issues as it ensures data integrity and preventing tampering [8]. In agri-food chain, blockchain technology enables the recording of transactional data or other digital events in its distributed database system by chain participants through decentralized consensus. The technology offers an accurate, decentralized, immutable and consensus process that promotes trust and transparency between stakeholders. There are three main types of blockchain which include public, private and permission networks (Hyperledger Fabric). The public network is more secured but too slow and expensive for application to livestock scenario. On the other hand, private networks are fast but are not suitable for livestock-traceability as they are not visible to the public. In permissioned networks, participation is control but visible to the public. Therefore, the permission blockchain is suitable for livestock-traceability due to their scalability, low latency and low cost of development [9]. Thus, this paper proposes Hyperledger Fabric for the poultry management system. The system is proposed to manage the history of poultry rearing from a hatchery, farm, processing plant to distribution and retail outlet for customers. The objective is to manage the entire history of poultry products. Thus, the main tasks of the system include the recording of the information associated with poultry rearing (from a hatchery to a grow out house), status report of the farm activities monthly, track the movement of product through the supply chain until delivery to the final consumer through a retail outlet.

2. Poultry, Human Health and Climate Change

Despite the benefit of livestock to human health and its future market demand projections, it still poses a threat to human health considering the intensity of the interaction between humans and birds. Globally, there is a constant transmission of disease between humans and animals on daily bases as 61% of the animal diseased are zoonotic [10]. Meanwhile, livestock are the major global contributors of antimicrobial resistance (AMR)-an emerging threat to human health-and is expected to rise by about 70% by 2030 [1].

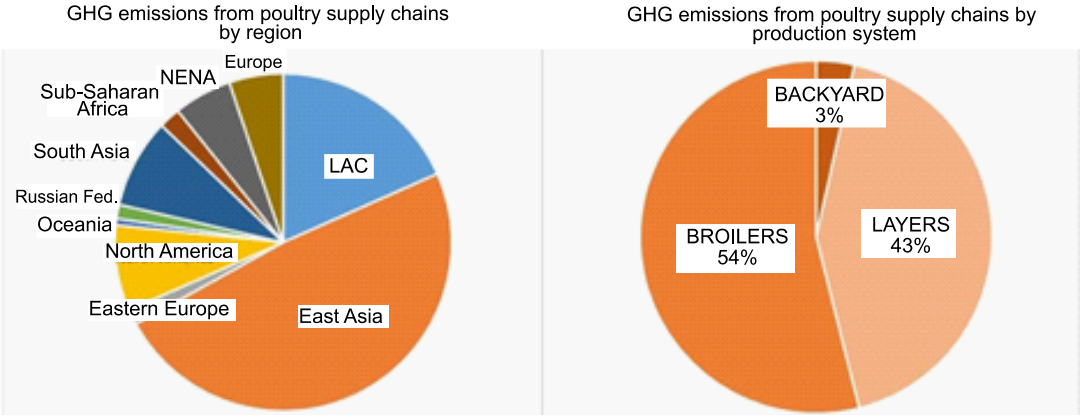


Fig. 1. Poultry supply chains contributions to Greenhouse gas emissions [1]

Since 1997, the majority of pandemics have been traced to have animal origin such as H5N1 avian influenza. The current COVID-19 pandemic is originated from animals. It was believed to be from bats and transmitted to humans through an unknown intermediate animal host [11]. The outbreak is now estimated to have infected over 34.8 million people and result in the death of over a million people globally as of October 2020 [12]. In the second quarter of 2020 alone, 14.0% of working hours, which is equivalent to 400 million full-time jobs, were lost relative to the last quarter of 2019. The COVID has paralyzed has affected almost every facet of human life [13]. In order to reduce the negative impact of livestock while increasing the positive contribution to human health, livestock information tracing through all the stages of production should be made a priority, particularly that of poultry. This will ensure food safety and fascilitate recall in case of contamination.

The direct or indirect emission of greenhouse gases (GHG) through manure management or feed production activities by the poultry sector contributes to climate change. Generally, the poultry supply chain constitute about 11% of the total GHG emission from the livestock-equivalent to about 836 million tons of CO2 [1]. As shown in **Fig. 1**, East Asia leads in the emission chart per region around the world-This suggests that decisive action is needed particularly in the East Asia. In course of poultry production, nutrient and pesticide pollution from the use of important land areas leads to biodiversity losses. In Brazil and Argentina, the expansion of arable land for feed production is a direct driver of deforestation [1, 14]. Also, the poultry sector results in air pollution by the emission of atmospheric ammonia, hydrogen sulphide and other odour particulate [15]. To address this climate issue pose by the poultry sector, accurate estimates are needed through systematic tracking of information in the sector. The integrity and reliability of the information are critical in establishing an accurate estimate in order to take the appropriate action.

3. Hyperledger Fabric For Poultry Traceability

Hyperledger Fabric is a permissioned blockchain network with a decentralized ledger that allows participating organizations to form a consortium. The organizations in poultry supply chain information systems include farmers, processors and distributors, with customers as the end of the chain. Each organization is recognized as members within the network and is saddled with the responsibility of creating network peers. Peers are issued identification credentials by Fabric Certificate Authority (CA). Transactions invocations are conducted by clients of their peers. As shown in the framework in **Fig. 2**, a client (mobile app, organization portal and web application) can interact with Hyperledger Fabric via REST web service or Hyperledger fabric SDK. The main components of the Hyperledger fabric network for poultry farms are discussed in the following subsection

.

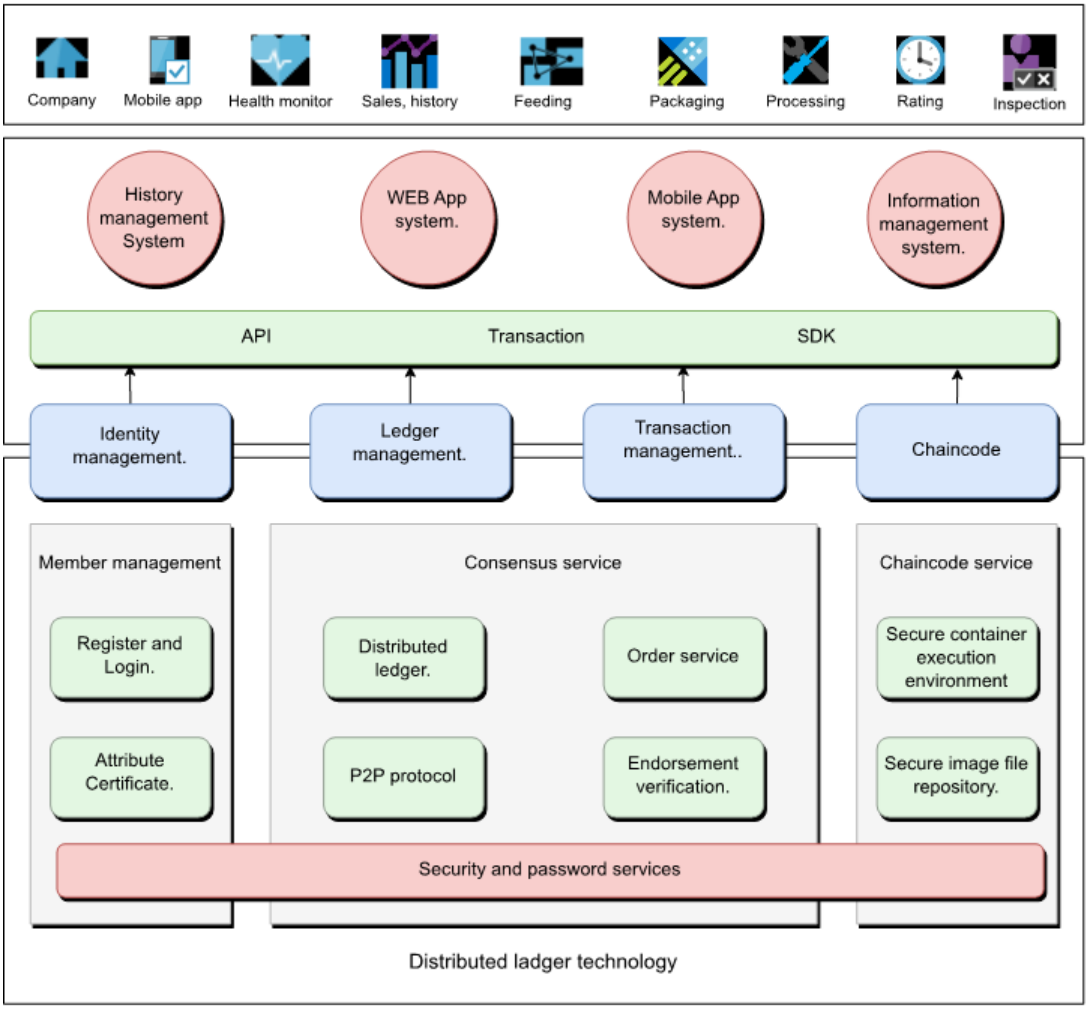


Fig. 2. Poultry product operational framework of Hyperledger-based traceability

3.1 Ledger

In this section, we examine how the Hyperledger-based traceability system organizes important objects in its ledger world state that reflect the poultry product supply chain. Each organization maintains data using objects as states within the ledger world state database. The business objects interactions such as query and update with the world state and blockchain are managed by smart contracts. The chaincode the smart contracts partition the ledger world state into different namespaces. Only the smart contracts in the same channel can share the same world state while the blockchain can contain transactions related to various chaincode world states.

3.2 Peers and smart contract

A blockchain network is comprised mainly of a set of peer nodes that host the ledgers and smart contracts. Transactions generated by smart contracts are immutably recorded into the ledger. In the poultry scenario, each organization consists of peers that house the copies of the ledgers and smart contracts. Each organization within the network consists of a set of peers that maintain the instance of the ledger. The peers in the network enable administrators and applications to interact with the Hyperledger service via a set of APIs. More so, the smart

contracts are simply business logic implementation in an executable program. Various rules, terms, concept definitions and processes that govern the poultry product supply chain are design here. In our scenario, each organization is empowered to upload information deem critical for tracking and tracing product history. In hyperledger fabric, roles are assigned to peers in the network which includes - Endorser peer, Anchor peer, Orderer peer and General peer. The endorser peer handles the reception of transaction invocation request, validation and execution of chaincode, the anchor peer is responsible for updating the orderer peer and broadcasting of transactions results to other peers in the network, while the orderer peer serves as the heart of the network that maintains consistency across the network.

4. Implementation of the Hyperledger Fabric-based Poultry Product Management System

The overall designed workflow between organizations through the channel in the system is as shown in Figure 2. From the figure, the organizations (Grower farm, Processing plant, and Distributor) are connected through a single channel with each participant (organization) having three peer nodes in total. In our design, each organization has two endorser nodes and one anchor peer. The endorser node holds a smart contract and validates transactions while the anchor peer serves to discover other peers from the connected organizations in the network. Moreover, the organization has admin and ordinary user roles.

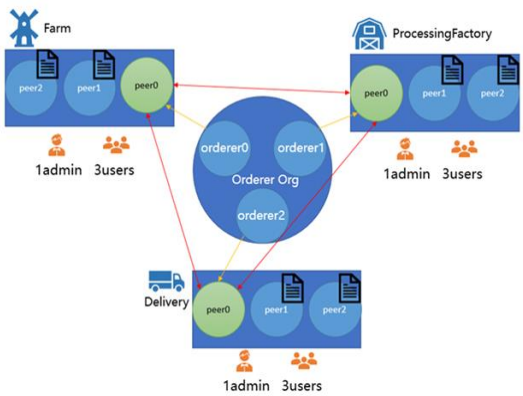


Fig. 3. Workflow between organizations in the Hyperledger fabric.

The system operation is initiated when a peer invokes a transaction such as a farmer documenting the feed and vaccination information into the system. The transactions are initiated using the Client App-Web App or Mobile App. The Endorser Peer receives a transaction from an associated Client App and applies a business logic chaincode for processing. If the result sent back to the Client App is approval, the app sends it to the Orderer Peer where the transaction is sent to Anchor Peers for onward accessibility to all General Peers.

5. Conclusions

In this paper, we highlighted the problems faced by the poultry sector while presenting the huge economic and health benefits it presents. We proposed the used of hyperledger fabric to address the problems of consumers' confidence, disease traceability and accurate estimate of the climate impact of the poultry sector. We discussed various aspects of the hyperledger and present a general architecture by identifying the organizations in the poultry product supply chain which include farms, processing plants and distributors. The hyperledger technology offers an accurate, decentralized, immutable and consensus process that promotes trust and transparency between stakeholders, and also facilitate easy identification of the sources of contamination in case of recall to safeguard public health. For feature works, we will focus on the implementation of the design, testing the software and analysing its effectiveness.

Acknowledgment

This paper has been written with the support of Jeollannam-do ('2019 R&D supporting program' operated by Jeonnam Technopark').

References

[1] A. Mottet and G. Tempio, "Global poultry production: current state and future outlook and challenges," *World's Poultry Science Journal*, vol. 73, no. 2, pp. 245-256, 2017/06/01 2017.

[2] E. Chowdhury and A. Morey, "Intelligent packaging for poultry industry," *Journal of Applied Poultry Research*, vol. 28, no. 4, pp. 791-800, 2019.

[3] N. Alexandratos and J. Bruinsma, "World agriculture towards 2030/2050: the 2012 revision," 2012.

[4] C. Heneghan. "Food safety top concern for consumers: survey." Accessed Aug. 2015, Available: https://www.fooddive.com/news/food-safety-top-concern-for-consumers-survey/407395/

[5] K. G. Grunert, "How changes in consumer behaviour and retailingaffect competence requirements for food producersand processors," *Economía Agraria y Recursos Naturales-Agricultural and Resource Economics,* vol. 6, no. 11, pp. 3-22, 2006.

[6] A. A. Tudoran, A. Fischer, J. van Trijp, K. G. Grunert, A. Krystallis, and L. Esbjerg, "Overview of Consumer Trends in Food Industry: Deliverable D2. 1," *Recapt*, 2012.

[7] A. M. Okrent and A. Kumcu, "US households' demand for convenience foods," 2016.

[8] G. Zhao, S. Liu, C. Lopez, H. Lu, S. Elgueta, H. Chen, and B. M. Boshkoska, "Blockchain technology in agri-food value chain management: A synthesis of applications, challenges and future research directions," *Computers in Industry*, vol. 109, pp. 83-99, 2019.

[9] M. Stefanova and M. Salampasis, "Farm-To-Fork Traceability: Blockchain meets Agri-food Supply Chain," 2019. Available: https://efita-org.eu/wp-content/uploads/2020/02/34.-efita1.pdf

[10] IFAH, "The Costs of Animal Disease," in "A report produced for the International Federation for Animal Health ", *Oxford Analytica* 2012, Available: https://healthforanimals.org/component/attachments/attachments.html?task=attachment&id=34.

[11] WHO. "World Health Organisation recommendations to reduce risk of transmission of emerging pathogens from animals to humans in live animal markets or animal product markets." *World Health Organisation*, 2020, Available: https://apps.who.int/iris/rest/bitstreams/1278909/retrieve

[12] WHO. "Coronavirus disease (COVID-19) Weekly Epidemiological Update and Weekly Operational Update." *World Health Organisation*, 2020, Oct. 10. Available: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20201005-weekly-epi-update-8.pdf

[13] ILO, "ILO Monitor: Covid-19 and world of work.," June, 30th 2020, Available: https://www.ilo.org/wcmsp5/groups/public/@dgreports/@dcomm/documents/briefingnote/wcms_749399.pdf.

[14] D. C. Morton, R. S. DeFries, Y. E. Shimabukuro, L. O. Anderson, E. Arai, F. del Bon Espirito-Santo, R. Freitas, and J. Morisette, "Cropland expansion changes deforestation dynamics in the southern Brazilian Amazon," *Proceedings of the National Academy of Sciences*, vol. 103, no. 39, pp. 14637-14641, 2006.

[15] P. Gerber, C. Opio, and H. Steinfeld, "Poultry production and the environment—a review," *Animal production and health division, Food and Agriculture Organization of the United Nations, Viale delle Terme di Caracalla*, vol. 153, 2007.

Prediction Modeling For Vision Screening Of Heavy Transport Vehicle (HTV) Drivers

Tayaba Jawed, Dr. Isma Farah Siddiqui*, Dr. Qasim Ali
Department of Software Engineering, Mehran University of Engineering & Technology, Pakistan
[E-mail: tayabajawed@gmail.com, isma.farah@faculty.muet.edu.pk]
*Corresponding author : Dr. Isma Farah Siddiqui

Abstract

This paper presents the prediction modeling of the data generated by the system "Vision Screening Of Heavy Transport Vehicle (HTV) Drivers" proposed by the Sindh Institute of Ophthalmology & Visual Sciences (SIOVS), Pakistan. It is significant as through this the accidental rates can be predicted and the data for accidental rates would be gathered from National Highway Authority (NHA) Pakistan. The proposed prediction model provides a possibility to predict the accident rates which most probably occurs due to the low vision of Heavy Transport Vehicle (HTV) Drivers. For this work, Data Mining Techniques are being applied on the datasets of patients that are required and are generated by the system proposed by SIOVS at M9 Motorway and dataset for the accidental rates as well generated by the National Highway Authority (NHA). Historical Datasets are being categorized into useful data sets by running queries in the database servers and then by applying predictive analytics models the information is being gathered from the raw datasets. After getting the results, it is being suggested that the HTV Drivers that are prone to accidental rates shouldn't be allowed to drive HTVs and they can only be allowed to get back to their job once they are properly being treated for the eye diseases they have which are relatively the cause of the accidents.

***Keywords*:** Vision prediction, Heavy Transport Vehicle, Health Informatics

1. Introduction

These days Health Information Technology (HIT) is very common for Information Processing which involves both computer software as well as computer hardware that deals with the storage, retrieval, sharing, and use of health care information, data, and knowledge for decision making. [1] Health / Eye Care Organizations nowadays use Health Management Information System (HMIS) that can be used to collect data of patients on daily basis that is required by the eye care program. [2] The Health Care Organizations / Institutes carries an enormous amount of raw data and for a particular purpose i.e. Decision Making, useful information could be collected from the raw data and to find excellent results these type of substantial amount of databases really require to be gouge out. To enhance productivity and quality of patient care, Information Technology offers health care industry significant potential.

With the extensive use of Health Care Information Systems commonly known as Electronic Health Records, there is a remarkable scope for improving the way healthcare is delivered by resorting to the power of big data [6]. For healthcare decision making data mining and predictive analytics are the most important tools. To deliver the improved healthcare services, the literature has reported strives for knowledge discovery from the big data to

however, there appears no attempt for evaluating and synthesizing the available information on how the big data phenomenon come up with the better outcomes for the delivery of excellent Health Care services.

There are many Eye Institutes that are using Hospital Management Information System (HMIS) to collect the eye patient's data. In collection with the points above mentioned, eye disease datasets have been elected from a system named as "Vision Screening in HTV Drivers at National Highway", implemented by Sindh Institute of Ophthalmology & Visual Sciences (SIOVS) Pakistan, this is the only autonomous Institute in its regional province that is providing the facility of Vision Screening for Heavy Transport Vehicle (HTV) drivers at M9 Motorway. Through this system everyday a large number of Heavy Transport Vehicle (HTV) drivers at M9 Motorway used to be examined by the Optometrists placed by SIOVS, on daily basis. SIOVS is playing this crucial role with the help of National Highway Authority (NHA) on M9 Motorway which is highly appreciated.

2. Literature Review

Buntin, M. B., Burke, M. F., Hoaglin, M. C., & Blumenthal, D. [1] To promote the assumption of Electronic Health Records and stimulus revolution in Health Care delivery, an endeavour effort is under way. To govern its effect on outcomes, the recent literature on health information technology has been reviewed, including efficiency, quality regulation, productivity and provider satisfaction. We came to know that 92% of the recent articles on Health Information Technology reached conclusions that were positive overall. We also found that the benefits of the technology are beginning to transpire in smaller practices and organizations, as well as in large organizations that were early begetter. However, there remains a problem among some providers it is still a dissatisfaction with Electronic Health Records and it's a barrier to attain the potential of Health Information Technology. This actuality call attention to the need for studies that document the challenging exposures of implementing Health Information Technology more specifically and how these challenges might be addressed.

Faal, H., Cook, C., Thulasiraj, R. [2] Health Information is one of the six components of a health care system. With the optimum aim of providing a better service and terminating preventable blindness, ee need to collect the information about each of the components in order to observe, plan and monitor. A well designed Health Management Information System (HMIS), whether at local, district, or national level it helps us to manage all the Information needed by (and generated by) the Eye Care programme. The Health Information Systems framework can be a useful thinking tool, whether we are planning a new Hospital Management Information System (HMIS) or assessing / evaluating an existing Hospital Management Information System (HMIS). It can help us to assist / determine that what questions the Hospital Management Information System (HMIS) should be proficient to answer (which in turn determines how we design it) and plan the implementation of our HMIS. Central to both of these is a recognition / consciousness of how our designed / planned Hospital Management Information System (HMIS) will integrate with any others in existence; or, where there are no others, how ours can be lengthen / extended to prop up other related Health Care Services.

Decina, L. E., & Staplin, L. [3] Those who were unenlightened of the fact that they are driving with weak eyesight, at the time of renewal of their driving license, it is examined that the Visual Examination of drivers were conducted. Every Driver's vision should be examined so that the accidental rates could be decreased.

Desapriya, E., [4] No doubt in driving Vision plays a significant role. To drive safely, the road signs, the traffic boards, the markings on the roads, the under construction part, the traffic cones and the vehicles moving on roads should be easily visible to the drivers. The vision of all the drivers should be examined to avoid accidents or any other kind of mishap while driving.

Homayounfar, P. [5] To extract / retrieve the knowledge from the raw data available in database of Health Care Institutes, Hospital Management Information System (HMIS) are being used from the past few years. That data would be used for fetching high quality results / knowledge and decision making and for getting the high-quality results.

3. Data Consolidation

3.1 Data Gathering

For this research, we have gathered the necessary raw data for Data Mining. Historical Data is gathered for getting the informative and useful data regarding patients and National Highway Authority (NHA) accidental rates. The data is further refined.

3.2 Extracting Useful Data

Useful Data is retrieved from the raw data by running the SQL Queries. Through SQL Queries raw data would be converted into the informative data on which Machine learning Techniques and Predication Modeling Techniques were applied to get the desired results.

Fig. 1. Methodology

4. Flow Chart

Fot this paper, historical data is gathered. Then the raw data underwent through pre-processing by removing the null values, missing values and duplicated values. For prediction modeling we applied the Data Mining Technique named "Data Cleaning and Preparation" on the pre-processed datasets. The flowchart in **Fig. 2** below shows the basic steps carriedout in this process.

Further, we applied Machine Learning Algorithm named "Two Class Decision Forest" on the data to get the desired reults. Through all this process we finally predicted that how many HTV Drivers are prone to accidents, which Eye Diseases are dangerous while driving and drivers with those Eye Diseases shouldn't be allowed to drive HTVs until and unless they are being treated.

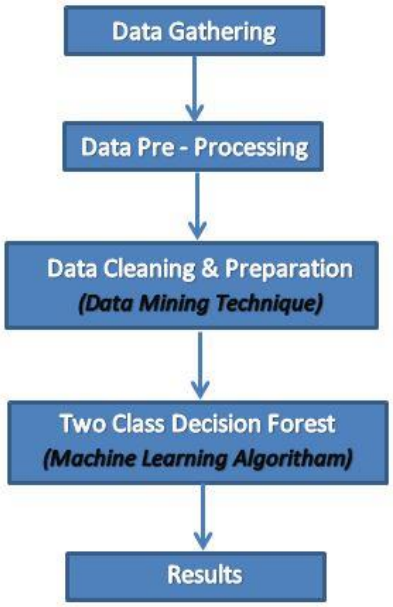


Fig. 2. Methodology Flowchart

5. Implementation

In this research, we have gathered data of patients collected from a Hospital named "Sindh Insitute of Ophthalmology & Visual Sciences (SIOVS)" this Eye Hopsital is located in Hydrabad, Sindh, Pakistan and we collected another dataset regarding Accidental rates fromNational highway Authority (NHA) Pakistan with their permission. Then, SQL Queries have been run on gathered data through SQL Server Query Mangaement 2018, data has been cleaned by running SQL Queries. We have run different queries for removing the null values and missing values. Now then we got the cleaned data further that data was coverted into the CSV File. Further, on that CSV Data File, the Machine Learning Algortitham have been applied named "Two Class Decision Forest" by using Microsoft Azure Machine Learning Studio. We have applied the "Two Class" Algoritham because the output labels are divided into two classes "Fit" and "Not Fit". Now, the End User can see the data in different forms like the End User can visualize the data in the form of Graphs and can genertae the Excel Sheet as well.

6. Evaluation

We have used different partitions of data and by using the partition of data, testing have been done through "Cross Out Evaluation". We have used the "Cross out Evaluator" to evaluate the Algorithm. Initially we have applied two different algorithms of Machine Learning named "Two Class Decision Forest" and "Two class Logistic Regression". Now, to find out which algorithm is deriving the accurate predictions, we further evaluated each machine learning algorithm for its precision and recall finding. We have evaluated to find out the best algorithm which can give the best results. The performance of machine learning algorithms have been evaluated by using hold out method (usually known as percentage split) and by using cross validation method. The evaluation and the results through "K-Fold Cross Validation" have been shown in the figure 3, 4, 5 and 6. As per our experiments with two algorithms, the Two Class Decision Forest is found to predict better with respect to its result of precision and recall.

6.1. Hold-out:

It is the simplest evaluation method where the data set is divided into two parts, one for training the model and the other one is to test the trained model. The purpose of this technique is to test the model on the data that is different from the data on which the model learned [8].

For this research, we performed a repeated hold-out evaluation where we evaluated the model using different data partitions for four times. The partitions we used are:

- 20-80: 20% data for training, 80% data for testing
- 40-60: 40% data for training, 60% data for testing
- 60-40: 60% data for training, 40% data for testing
- 80-20: 80% data for training, 20% data for testing

We observed that the best result is achieved through the fourth case (i.e. 80% data for training and 20% data for testing).

6.2. K-Fold Cross Validation:

It is an evaluation technique that evaluates the machine learning models on limited data using resampling procedure. The procedure requires to specify the value of "k", where k specifies the number of groups in which the available data should be divided for evaluation. Such evaluation is called k-fold cross validation [24].

For this research we used 10-fold cross validation for evaluating the performance of machine learning models used. Mean and Standard of Precision, Recall, Accuracy and Average Log Loss of 10-Folds values have been shown in the table. The results from this evaluation are shown in **Fig. 3**.

DECISION FOREST				
K-Fold	**Accuracy**	**Precision**	**Recall**	**Average Log Loss**
Mean	0.89	0.92619	0.914286	1.639383
Standard Deviation	0.159513	0.182522	0.153567	3.285601
LOGISTIC REGRESSION				
K-Fold	**Accuracy**	**Precision**	**Recall**	**Average Log Loss**
Mean	0.88	0.909524	0.914286	0.395614
Standard Deviation	0.154919	0.182643	0.153567	0.363152

Fig. 3. Results from Evaluation

Since both evaluation techniques marked One-vs-all multiclass method as the best one, we convert this model into a web-service that can be used in the intended application for making the predictions. Furthermore, we prefer to convert the model created with cross validation technique over model created with hold-out evaluation, into the web service, for the following reasons:

- Cross validation doesn't require multiple iterations like hold-out approach for accurate predictions.
- It reduces the variance of estimates.
- It can work on limited data set.
- Trains and validates the model in single turn.
- Last but not the least, it evaluates the model as well as data.

The results of these experiments are shown in the graphs in **Fig. 4** to **Fig. 7**.

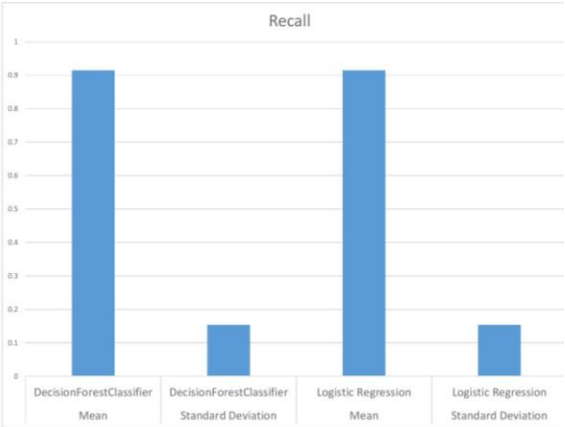


Fig. 4. Recall Result of the two ML Algorithms

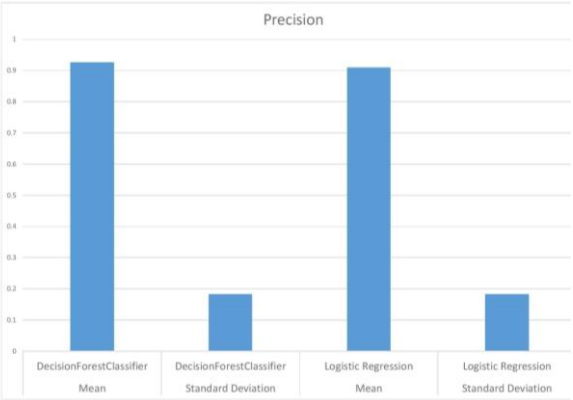
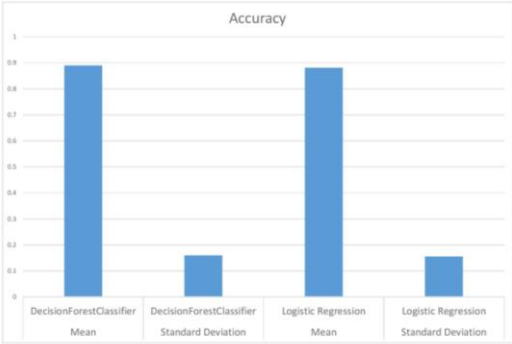


Fig. 5. Precision Result of the two ML Algorithms

7. Scope and Limitation

As the large data set would be required to apply the Data Mining Algoritham. The performance of the system should be very good with good storage and processor for fast perfomarnce otherwise performace will be degraded because of large dats sets. The target audience for this



research research is the Heavy Transport Vehicle (HTV) Drivers only.

Fig. 6. Accuracy Result of the two ML Algorithms

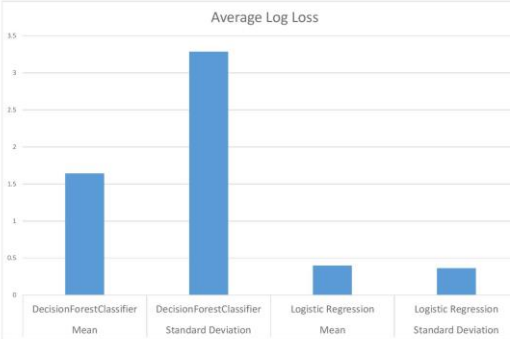


Fig. 7. Average Log Loss of the two ML Algorithms

In future, we will extend our experiments with other advanced machine learning algorithms for generating more better prediction model.

9. Conclusion

In this paper, we present an interactive method to manipulate a raw historical dataset into informative dataset by applying prediction modeling technique. To predict that the patients that are prone to Eye Disease could be the reason of increased data and those drivers who are driving Heavy Transport Vehicles (HTVs) with Eye Diseases should be first treated in any renowned Eye Hospital and then all of those drivers should be allowed to drive the HTVs.

References

[1] Buntin, M. B., Burke, M. F., Hoaglin, M. C., & Blumenthal, D. (2011). The benefits of health information technology: a review of the recent literature shows predominantly positive results. Health affairs, 30(3), 464-471.

[2] Faal, H., Cook, C., Thulasiraj, R. (2010). Managing information in eye care programmes: the health systems perspective. Community Eye Health, 23(74), 50-52.

[3] Decina, L. E., & Staplin, L. (1993). Retrospective evaluation of alternative vision screening criteria for older and younger drivers. Accident Analysis & Prevention, 25(3), 267-275.

[4] Desapriya, E., Harjee, R., Brubacher, J., Chan, H., Hewapathirane, D. S., Subzwari, S., & Pike, I. (2014). Vision screening of older drivers for preventing road traffic injuries and fatalities. Cochrane database of systematic reviews, (2).

[5] Panella, M., Marchisio, S., Stanislao, F.: Reducing clinical variations with clinical pathways: do pathways work? Int. J. for Quality in Health Care 15(6), 509–521 (2003).

[6] Douglas, G. P., Deula, R. A., & Connor, S. E. (2003). The Lilongwe Central Hospital Patient Management Information System: a success in computer-based order entry where one might least expect it. In AMIA Annual Symposium Proceedings (Vol. 2003, p. 833). American Medical Informatics Association.

[7] Holdout Evaluation. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA, 2011.

[8] Daniel Berrar, in Encyclopedia of Bioinformatics and Computational Biology, 2019

Interpretability of Machine Learning Based Forecasting Models of Wind Power Generation in Jeju Island, South Korea

Young-Suk Han[1], Donghyun Park[2], Minwoo Baek[1] and Jae-Yoon Jung[1,2]

[1]Department of Industrial and Management Systems Engineering, Kyung Hee University, South Korea
[2]Department of Big Data Analytics, Kyung Hee University, South Korea
[e-mail: ysjhan@khu.ac.kr, pdh@khu.ac.kr, bjh713@khu.ac.kr, jyjung@khu.ac.kr]
*Corresponding author: Jae-Yoon Jung

Abstract

Under the 3020 plan for implementing renewable energy in Korea, the proportion of generating renewable energy is increasing, and over 95 percent of new power generation facilities are focused on the distribution of renewable energy, such as solar and wind power. In particular, wind power generation is used as a base power in actual system operation in Jeju, South Korea, and therefore accurate power generation forecasting is necessary to maintain power supply stability and reduce power production losses. In this research, machine learning techniques are used to predict wind power generation through weather forecast information and past wind power generation trends. Additionally, by using machine learning models along with machine learning interpretation methods, main variables can be identified and predicted results can be analyzed. To predict the exact amount of power generation in Jeju Island, we predict wind power generation of Hankyung Power Plant in Jeju Island, obtain knowledge of wind power generation prediction by using explaining machine learning models, and trust the prediction of the trained machine learning model.

***Keywords*:** Wind power generation forecasting, machine learning, machine learning interpretability

1. Introduction

Recently, wind power has emerged as one of important renewable energy sources together with solar power. However, since wind power is problematic since it depends on uncontrolled weather conditions, it is necessary to predict wind power based on weather conditions to sustain the stability and reliability of power generation [1-3].

In the case of Jeju Island, Korea, wind power is used as a base power in acutual system operation, and so accurate power generation predictions are needed in the real-time power trading market to maintain power supply stability and reduce power production losses.

Explaining machine learning models is a way of obtaining knowledge through learning new facts and gathering information. It can also help verify the predictions of models, enhance the model, and attain new insights into the problems that needs to be solved [4].

Based on the analysis results by using the machine learning model that explains and predicts wind power generation of Hankyung Power Plant in Jeju Island in this study, we obtain knowledge regarding wind power generator prediction, and gain confidence in the prediction of the machine learning model.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2017H1D8A2031138).

2. Wind Power Generation Prediction and Interpretation

2.1. Machine Learning Modeling and Evaluation

Three types of models such as linear regression, nonlinear regression, and decision tree and ensembles, are studied to build wind power prediction models. We used linear regression (LR) and partial least squares (PLS) models for the linear regression type, support vector machines (SVM) for the nonlinear regression type, and decision tree (CART), random forest (RF), bagging regression (Bagging), adaptive boosting (ADA), and extreme gradient boosting (XGB) for the decision tree and ensemble type.

Two evaluation measures are used to compare the predicted performance among the machine learning models: root mean square error (RMSE) and R -squared.

2.2. Machine Learning Interpretation

By interpreting a black box model with an interpretability method, we identify and comprehend the relationship between wind power generation volume and weather forecast. The data were collected at Hankyung Power Plant in Jeju Island.

This study uses the global model interpretation methods, which enables the understanding of the model's effect in the whole range of variables. In particular, using the methods such as feature permutation importances and partial dependence plot, the application result of the global interpretability method to the machine learning model explains the prediction of wind power generation in Jeju Island.

The feature permutation importance is a technique of determining the importance of features by permuting the value of features and measuring model prediction errors. It is easy to interpret and provide a global insight regarding the model behavior.

The partial dependence plot (PDP) is a technique of showing the average relationship between one or more input variables and the prediction of the black box model. PDP depicts the marginal effect on features that influence the prediction.

3. Experiments

3.1 Dataset

We collected a total of four years of wind power and weatehr forecast data from 2014 to 2017 to predict wind power generation in Jeju.

Wind power was measured at 10-minute intervals. Wind power generation data were collected from wind power plants in the Hankyung region of Jeju Island. In the case of wind power plants in Hankyung region.

The goal of this study is to predict average wind power generation per hour after four hours. **Fig. 1** describes in detail the temporal relationship among the weather forecast, the average wind power generation in the past, and the average wind power generation in the future (after 4 hours) to be predicted. The weather forecast data were also collected around the west side of Jeju Island, in which the Hankyung power plant is located. The weather forecast is announced every three hours from 2 a.m. every day. The forecast target time (t+4) is four hours later the forecast time (t). Additional variables include a total of 28 variables such as the number of turbines at target

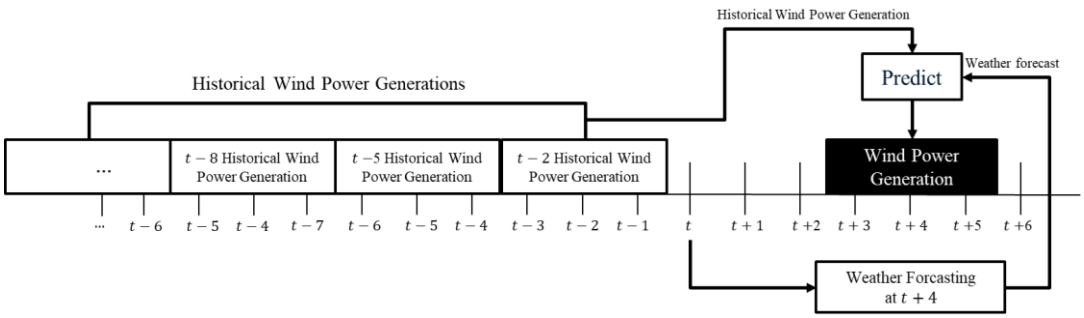


Fig. 1. Temporal description of the scenario of wind power generation forecasting based on weather forecast

forecast time and information about target time (e.g., season, month, day, and timezone). Both wind power data and weather forecast data were extracted for a train set from 2014 to 2016 and for a test set from 2017.

3.2. Machine Learning Modeling

A total of eight machine learning models were trained and tested. The performance comparison is illustrated in **Table 1**.

Table 1. Performance comparison among ML-based forecast models

Type	Method	R-squared	RMSE
Linear model	LR	0.783	623.651
	PLS	0.743	678.475
Nonlinear model	SVR	0.761	654.033
Decision tree and ensemble model	CART	0.594	852.512
	Bagging	0.792	609.850
	GBR	0.802	596.244
	ADA	0.760	655.741
	XGB	**0.806**	**589.784**

As summarized in **Table 1**, XGBoost had the highest performance among linear regression models such as LR and PLS, non-linear regression models such ass SVR, and ensemble models such as CART, Bagging, GBR, ADA, and XGB. Therefore, we will try to explain the XGBoost model in detail in the next subsection.

3.3. Machine Learning Interpretability

3.3.1. Features Permutation Importance

The feature permutation importance method was used to determine the importance of features that are used in the wind power forecasting model in Jeju Island. **Table 2** shows the top-6 important features of the wind power forecast model based on XGB.

Table 2. Feature permutation importance

Rank	Feature	Weight
1	n_turbine_in_use	0.4997 ± 0.0076
2	y_t2h	0.2451 ± 0.0073
3	wind_speed	0.0692 ± 0.0029
4	y_t5h	0.0260 ± 0.0011
5	wind_relationship	0.0228 ± 0.0015
6	temperature	0.0161 ± 0.0023

3.3.2. Partial Dependence Plots (PDP)

The partial dependence plot was adopted for showing how specific features affect wind power generation. Since it is difficult to show all features, we select and show features that significantly affect the prediction of wind power generation. In order to select features that greatly affect wind power prediction, the PDP was drawn based on the top-6 important features of XGB as shown in **Fig. 2**.

In the table, the x-axis of the plot is the value of the feature, and the y-axis of the plot is the amount of wind power generation forecasting. The bar below the plot represents the distribution of data for each interval. In the case of n_turbine_in_use, wind_speed, y_t2h, and y_t5h feature, it could be intuitively predicted that the amount of wind power would increase as the value increases, which can be visually confirmed through PDP. In the case of temperature feature, it has influence on increasing the wind power generation when it is less than 10.5 degrees, and has small influence on increasing the wind power generation when it is greater than 10.5 degrees.

4. Conclusions

This study presented the wind power generation prediction model based on weather forecast information and wind power generation trend in the past. Among 11 machine learning techniques for wind power forecasting in Hankyung wind power plant in Jeju Island, South Korea. Among them, XGB showed the highest performance in terms of R-squared and RMSE.

The top-6 important features of XGB in forecasting wind power generation were the nubmer of turbines in use, three weather-related variables (wind power, temperature, wind relationship) and two historical wind power generation variables (2 hours and 5 hours before).

PDP confirmed that the features that can be intuitively predicted to increase wind power as the value increases actually operate in the model. and the increase in wind power generation can be noticed when the temperature feature is below 10.5 degrees.

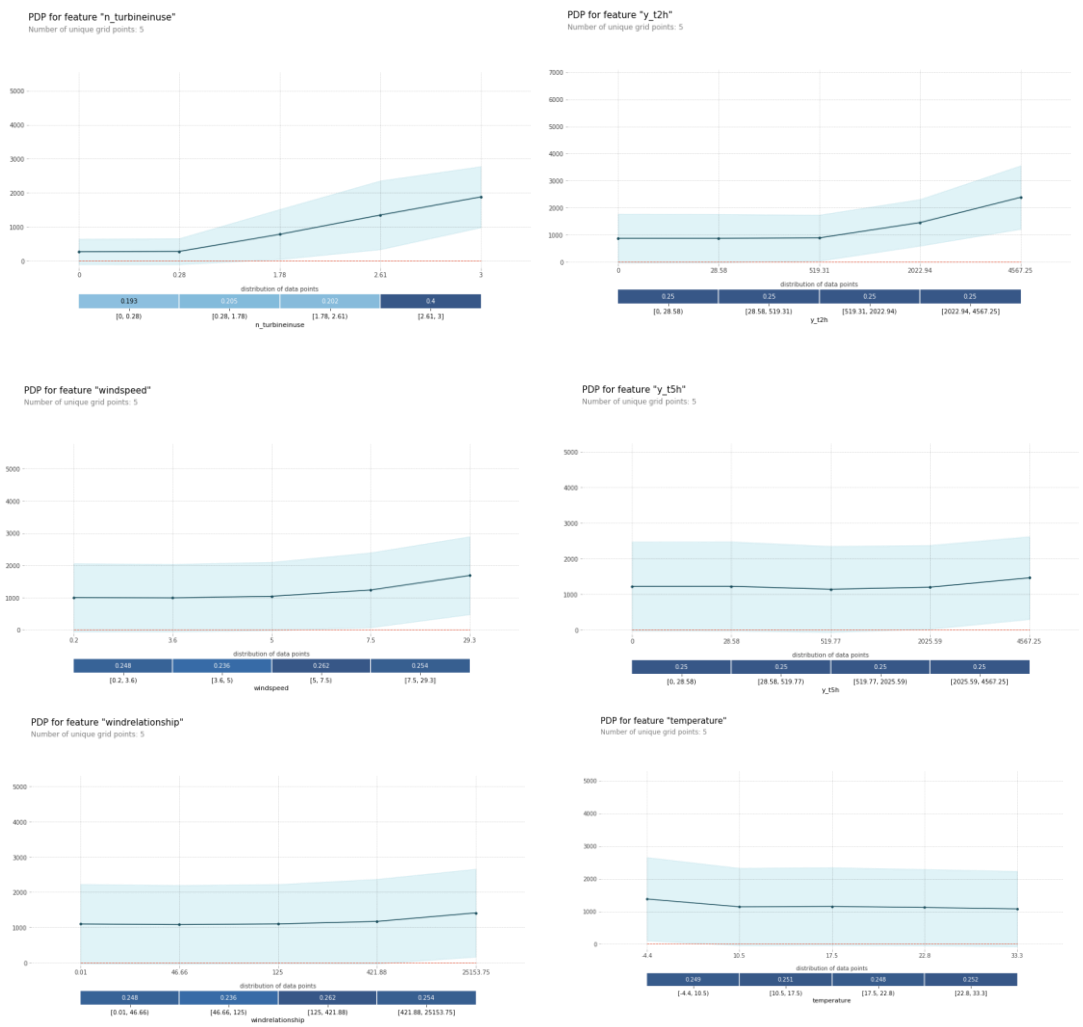


Fig. 2. The partial dependence plots (PDP) of six important features in XGB models

References

[1] Traber, T. and Kemfert, C. "Gone with the wind? Electricity market prices and incentives to invest in thermal power plants under increasing wind energy supply," *Energy Economics*, Vol. 33, No. 2, pp. 249-256, 2011.

[2] Yurong Zhang, et al.,"Unit commitment considering effect of load and wind power uncertainty,"IEEE Workshop on Advanced Research and Technology in Industry Applications (WARTIA), pp.1324-1328, 2014.

[3] S. Li, D. C. Wunsch, E. A. O'Hair and M. G. Giesselmann "Using neural networks to estimate wind turbine power generation,"IEEE Transactions on Energy Conversion, vol. 16, no. 3, pp.276 –282, 2001.

[4] Adadi, A. and Berrada, M., "Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI)," IEEE Access, Vol. 6, pp. 52138-52160, 2018.

A Web-based AI Platform for Interactive Supercomputing Service

Guohua Li, Joon Woo, Seongjun Kim, Minwoo Kwon, Taeyoung Hong
Korea Institute of Science and Technology Information
Daejeon, South Korea
[e-mail: ghlee@kisti.re.kr]
*Corresponding author: Guohua Li

Abstract

In recent times, researchers in the field of computational science using supercomputing resources have initiated applying AI techniques. The demand for high-performance AI infrastructure and development platforms has been increasing. To address this, the national supercomputing infrastructure center of the Korea Institute of Science and Technology has constructed a high-performance graphics processing unit-based system – Neuron, to provide high-performance AI infrastructure environment service. However, traditional supercomputing service platforms based on high-performance AI infrastructure do not enable users to analyze and visualize data interactively. To address these limitations, we provide interactive supercomputing services by developing a web-based AI platform based on Jupyter, the most used development tool for the development environment. In this study, we use this platform integrated with lightweight directory access protocol (LDAP), one-time password (OTP), and Slurm using JupyterHub that can interconnect with a multi-user task environment, existing batch scheduler, and several functions to enhance security authentication. Users can use Jupyter notebook/JupyterLab on a supercomputing resource by selecting a resource type classified as a profile on the web. In addition, to address the problem of having to pay expensive bills, which occurs when users submit interactive jobs, we created a free and dedicated Jupyter queue to enable users to use resources in a cost-effective and efficient manner. Developing an IPyParallel plugin enables parallel programming in Python in a supercomputing environment with Jupyter notebook. This increased the use of supercomputing resources.

***Keywords*:** Jupyter, interactive supercomputing, AI platform, Python, parallel programming

1. Introduction

In recent times, there has been an increase in the demand for graphics processing unit (GPU)-based AI infrastructure in the field of computational science. This demand cannot be adequately managed using the existing resources as they are limited in number. Hence, a customized AI service environment that can access and utilize several platforms for AI-based research is required. Our study was conducted at the national supercomputing infrastructure center of the Korea Institute of Science and Technology (KISTI). This center operates on Nurion and Neuron, the No.5 Supercomputer. Nurion is a CPU-based system; Neuron is a GPU-based system that addresses the limitations of Nurion. Neuron started its official operation in May 2019 with a share-file system of Nurion, 141 GPU devices installed on 73 computation nodes, and a theoretical performance of 1.02 petaflops.

This study intends to provide insights on services to users using GPU-based Neuron as an AI platform. The use of traditional high-performance computing (HPC) service

This study was performed as a subproject of the KISTI project entitled "The National Supercomputing Infrastructure Construction and Service [K-20-L02-C01-S01]".

platform causes limitations in interactively analyzing and visualizing research data. To address these problems, we offer HPC services by developing a web-based AI platform using Jupyter, a development tool that users use frequently. Overseas national supercomputing centers also offer web-based interactive Jupyter services suitable for their system environment. In this study, we developed an AI service platform suitable for Neuron systems by analyzing case studies and the pros and cons of overseas research.

2. Related Works

2.1 NERSC Jupyter Service

The national energy research scientific computing center (NERSC) in the United States mentioned that it will provide a platform for large-scale scientific deep learning applications. Perlmutter, a next-generation system for machine learning, is scheduled to be introduced in 2021 [1]. The existing GPU-based Cori/Spin supercomputer environment for machine learning is available through the web-based Jupyter [2]. The Jupyter service enables users to interactively develop machine learning/deep learning, visualize, and analyze data in a high-performance infrastructure using supercomputing resources. As shown in **Fig. 1**, users can access the JupyterHub through a web browser, authenticate, allocate resources through Slurm scheduler, and deploy Jupyter notebook/JupyterLab. The computed nodes to be used interactively are allocated in the Cori supercomputer and a multi-node AI execution environment is configured with an IPyParallel cluster to provide a parallel programming environment in Python.

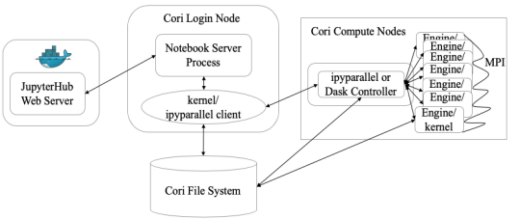


Fig. 1. NERSC Jupyter architecture

2.2 CSCS Jupyter Service

The Swiss national supercomputing center (CSCS) provides Jupyter services using the Piz Daint supercomputer. The services enable MPI in single/multimode through IPyParallel and mpi4py as shown in **Fig. 2**. In the CSCS architecture, although the notebook server process does not operate on a separate login node, they operate on compute nodes with other IPyParallel clients and engines. CSCS announced that it aims to provide Jupyter service in the cloud environment using Shifter container [3].

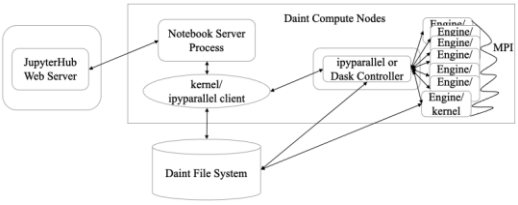


Fig. 2. CSCS Jupyter architecture

3. Platform Implementation

3.1 Service Components

To overlay the virtual 3D models on the augmented reality system, a marker in the input image was initially detected. We provided Neuron Jupyter services through a web site [4]. The services consist of Jupyter notebook, JupyterLab, Tensorboard, and IPyParallel. When a user attempts to login with registered supercomputer accounts, they can select supercomputing resources via a queue (resource templates) and execute the described Jupyter services. We offer the services by partitioning it into a free dedicated queue for environment installation, preprocessing, debugging, and other billing queues for applying models such as deep learning/machine learning. Billing is calculated based on the GPU usage rate.

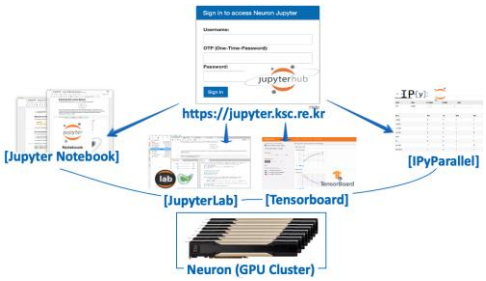


Fig. 3. Service components

3.2 System Components

We used *ldap-jupyterhub* as shown in **Fig. 4** to integrate LDAP and OTP authentication systems and developed a package named *kisti-batchspawner* (as shown in **Fig. 4**) that can connect with the Slurm scheduler responsible for resource provisioning. These integration packages were developed based on the open source JupyterHub. We integrated Tensorboard with Jupyter notebook and JupyterLab as a plugin. In addition, we integrated these packages with Slurm to use IPyParallel in Jupyter notebook, and developed functions to submit parallel jobs with IPython through a web interface (the *kisti-scheduler* shown in **Fig. 4**).

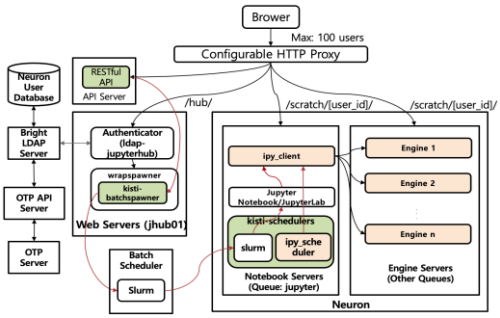
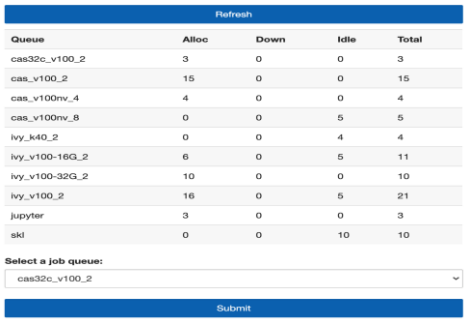


Fig. 4. System components

To visualize the status of supercomputing resource usage through the web as shown in **Fig. 5**, *RESTfull API* was developed for each feature.



Refresh

Queue	Alloc	Down	Idle	Total
cas32c_v100_2	3	0	0	3
cas_v100_2	15	0	0	15
cas_v100nv_4	4	0	0	4
cas_v100nv_8	0	0	5	5
ivy_k40_2	0	0	4	4
ivy_v100-16G_2	6	0	5	11
ivy_v100-32G_2	10	0	0	10
ivy_v100_2	16	0	5	21
jupyter	3	0	0	3
skl	0	0	10	10

Select a job queue:

cas32c_v100_2

Submit

Fig. 5. Resource status

3.3 Platform Deployment Environment

The deployment of Jupyter platform environment was executed based on the Anaconda virtual environment as shown in **Fig. 6**. Users can configure the environment through the automated installation scripts. The *kisti_conda_jupyter.sh* script automatically creates a virtual environment *notebook* and installs libraries such as *jupyterhub*, *jupyter notebook*, *jupyterlab*, *tensorflow*, and *cudnn*, that are required to be executed on the computing nodes in the environment. The *kisti_conda_plugins.sh* script installs *tensorboard* and *IPyParallel* plugins in the created *notebook* environment. It also includes the IPyParallel plugin developed by KISTI (shown in **Fig. 7**) for Neuron Jupyter.

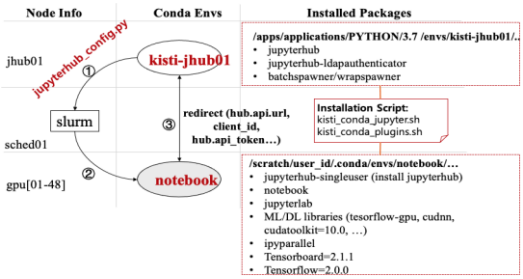


Fig. 6. Neuron Jupyter deployment environment

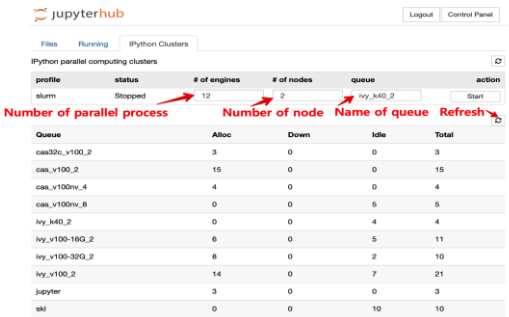


Fig. 7. Jupyter notebook IPyParallel plugin

4. Conclusion

This study was conducted with an aim of providing services by constructing a platform to enable a developed AI environment to be used as a web-based interactive supercomputing service. Users can use high-performance resources based on the web to develop applications in deep learning/machine learning environments. We launched the service through the national supercomputing center, and we aim to examine the platform using research based on computational science.

References

[1] J. M. Brandt, "HPC monitoring & analysis+ power 9 specifics," No.SAND2019-9734PE.

Sandia National Lab.(SNL-CA), Livermore, CA (United States), 2019.
[2] T. Rollin, "Jupyter at NERSC redefining the interface to HPC," *NERSC User Group Meeting*, July 2019.
[3] A. Sadaf, "Interactive supercomputing at the crossroads of HPC and cloud technologies," *GTC Europe*, October 2018.
[4] Neuron Jupyter, https://jupyter.ksc.re.kr

To Improve Software Testing in Crowd-based Outsourcing Environment by Relating Levels of Testing with Personality Characteristics

Zainab Umair Kamangar, Dr. Isma Farah Siddiqui*, Dr. Qasim Ali, Umair Ayaz Kamangar and Sunbul Sajid Khowaja
Department of Software Engineering, Mehran University of Engineering and Technology, Jamshoro, Pakistan
[e-mail: zainabumair11sw38@gmail.com, isma.farah@faculty.muet.edu.pk]
*Corresponding Author

Abstract

Crowd-based outsourcing is an emerging trend in testing, which integrates advantages of crowd-based outsourcing in software testing. Open call format is used to accomplish various tasks on network involving different types of testing levels and techniques at various places by software testers. This process makes software testing more reliable and error free, as it performed by multiple testers with different levels using different platforms and techniques. Suitable task performance by required skilled tester is a hard to do in crowd-based testing environment and also overload the platform to evaluate all the tasks submitted by testers. Crowdsourced software testing can lead to improper testing process as if it is not allocated the right task to right crowd with required skills and also not choosing the right crowd can lead to huge amount of results, which become difficult and time consuming for crowdsource manager for the identification of improper one. The main purpose of this research paper is to make crowd-based outsourced software testing more effective and reliable by relating association between the software tester, personality characteristic and different levels of software testing i.e., unit testing, integration testing and system testing in order to find appropriate personality characteristic for required testing level. This research paper has shown an observed experiment to determine which level of software testing (unit, integration and system level) suits which personality characteristic tester in crowd-based outsourced software testing environment. A total of 43 software testers from different software houses and firms in Pakistan were registered in this research to perform tasks at different software testing levels. Myers-Briggs type indicator (MBTI) test is used to identify the personality characteristic of each tester involved in this research study.

Keywords Crowd-based outsourcing, MBTI, Testing levels, Personality characteristics.

1. Introduction

Crowdsourcing is emergent technique based on distributed problem-solving model, which combinedly use human as well as machine computation. In 2006 the term ‘crowdsourcing’ was combinedly devised by Howe and Robinson. It is defined as crowdsourcing is an open call format which is used by organization to outsource the workers in participation on network with undefined, online labor [1].

Software testing is an analyzing process to evaluate features of software element and to distinguish the alterations between current and required circumstances (i.e., defects) of software elements [2]. Crowd-based outsourced Software Testing is the grouping of crowd-based outsourcing and software testing to use software testing in crowd-based environment (open call format) to get advantage of recruiting globally online connected software testers. These software testers prepare the test cases for software under consideration, send them to project manager of an organization for approval and perform testing if their test cases get approved from organization but the issue is that testing can be performed at different levels and it is must for manager to know which tester perform better at which level of software testing. There is no such related research is available which can relate the software testing level with crowd-based outsourced testers for selecting suitable testers which suits to relatable software testing level. Therefore, to resolve this issue an experimental research is performed, which shows that by selecting suitable testers for taking part in crowd-based outsourced testing and enable manager to choose appropriate software tester to increase the efficiency of this process of testing task assignment.

2. Related Work

The relationship between Crowdsourced software testing techniques and personality type has been proposed, it has identified that personality type with extrovert characteristic software testers are suitable at performing tasks related to black-box testing techniques while personality type with introvert characteristic software testers are suitable for performing tasks related to white-box testing techniques, and also it has been identified that software testers with personality type having dimension J are more suitable for performing tasks related to black-box testing [3].

The method for the task assignment in crowd sourcing environment for collaborative development has been proposed. It focuses on three factors to achieve goals which are worker time, capacity and task module complexity [4].

Performance can be affected by ability of tester and difficulty of testing task in crowd-based outsourcing. It is also found that functional testing task is performed improved when it is performed by skilled crowd tester as compared to one who have not got experience in crowd-based outsourcing environment, this difference between expert and beginner is clearer when task is more complex [5].

The current research of crowd-based outsourcing has been concentrating upon different areas, like; how crowd-based outsourcing is advantageous how to assign tasks of crowdsourcing and quality deliverables [6]. It has also been declared that a task cannot be performed satisfactorily if it is not allocated to desirable personality type worker, and also this statement raises new challenges for crowdsourcing tasks and involves a detailed consideration of allocating the workers with suitable tasks related to their personality type [7].

It has also been mentioned that, probability of task to become successful is increased, if task is assigned to appropriate personality type [8]. Task Assignment Model based on crowd-based outsourcing for software development has also been proposed [9]. The association between personality characteristic and task selection is identified and it also establishes association of personality types with both prize money and deadlines, some Personality characteristics were more attracted in choosing prize money than time limit while some were more interested to choose tasks related to the timeline [10].

The individual personality type, which is combination of four dimensions, as shown in **Table 1**, is classified based on the MBTI (Myers-Briggs Type Indicator) test, so there are 16 possible personality combinations, as shown in **Table 2**. To evaluate the personality characteristic of a crowd-based outsourced software development worker MBTI test is used as an instrument, as it identifies the similarities between tasks according to personality characteristic [11].

Table 1. MBTI Personality Characteristic's Dimensions

Extroversion (E)	**Introversion (I)**
Sensing (S)	**Intuition (N)**
Feeling (F)	**Thinking (T)**
Percieving (P)	**Judging (J)**

Table 2. MBTI 16 Personality Characteristics

ISTJ	**ISFJ**	**INTJ**	**INFJ**
ISTP	**ISFP**	**INTP**	**INFP**
ESTJ	**ESFJ**	**ENTJ**	**ENFJ**
ESTP	**ESFP**	**ENTP**	**ENFP**

The emphasis of this research paper is to determine which level of software testing (unit, integration and system level) is suitable for which personality type tester in crowdsourced software testing environment.

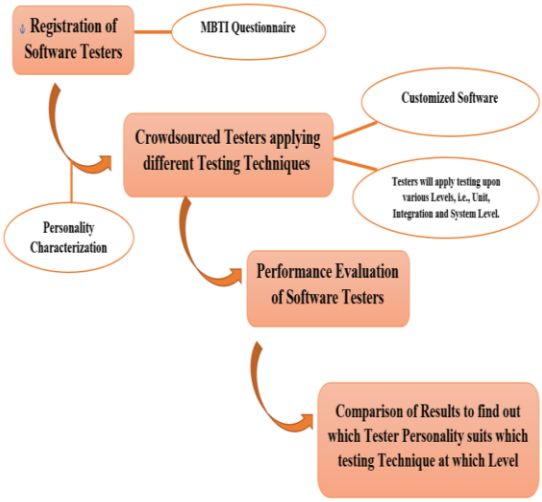


Fig. 1. Research Methodology

3. Methodology

To complete the core objective of this research, a tentative tactic was applied to tester's data set. Software testers from various firms and software houses in Pakistan, were involved in this research.

Firstly, 43 testers were registered for this research. Testers were required to fill the registration form which also contains a questionnaire on the basis of which their personality characteristic was identified.

The distinct personality characteristics are classified based on the MBTI (Myers-Briggs Type Indicator) test. In this study the personality of a crowd tester is evaluated by integrating MBTI as an instrument to identify personality type, as it is extensively used for similar assessment tasks based on personality type.

The study was conducted by providing testers with a costomized software program, its code and tasks to do. Testers were asked to complete thoes tasks on provided software to accomplish unit, integration and system level software testing aspects.

Customized software was also developed to set the tasks to accomplish the tasks related to levels to satify the study reqirements. Each tester have to do each and every task on the given software program.

4. Results and Discussion

In this study to accomplish the research experiment, the study was conducted on a customized software program, and setting up of defined tasks of each level of testing, also while registering the personality characteristic of each tester was identified. **Table 3** shows the personality characteristic and the consecutive number of testers registered of that personality characteristic.

Table 3. Personality Characteristic and number of testers registered having corresponding personality characteristic

Personality Characteristic No:	Personality Characteristic	Frequency
1	ISTP	2
2	ISFP	4
3	INFP	3
4	INTP	2
5	ISTJ	3
6	ISFJ	1
7	INFJ	2
8	INTJ	5
9	ESTP	3
10	ESFP	5
11	ENFP	3
12	ENTP	1
13	ESTJ	3
14	ESFJ	2
15	ENFJ	3
16	ENTJ	1

It is declared in the research methodology that testing tasks related to different software testing levels were performed by testers on customized software program and determined that testers who have introvert personality characteristic are appropriate for unit testing level as compared to extrovert personality characteristic, while the testers, who have extrovert personality characteristic are suitable in performing tasks related to integration and system testing level. Further it is also analyzed that testers with personality characteristics having J (Judging) dimension are better at executing tasks related to System testing level. **Table 4** indicates which personality characteristic execute tasks of which software testing level better based on scale from 1 to 10.

Table 4. Personality Characteristic and consecutive score average of testers in different testing levels based on scale (1-10)

PERSONA-LITY CHARAC-TERISTIC NO:	PERSONA- LITY CHARAC-TERISTIC	SOFTWARE TESTING LEVELS		
		Unit Testing	Integration Testing	System Testing
1	ISTP	9	5	2
2	ISFP	7	4	3
3	INFP	8	4	2
4	INTP	6	3	2
5	ISTJ	8	5	4
6	ISFJ	7	2	5
7	INFJ	8	4	4
8	INTJ	9	5	5
9	ESTP	3	9	5
10	ESFP	2	7	6
11	ENFP	4	8	6
12	ENTP	5	6	7
13	ESTJ	3	7	9
14	ESFJ	2	9	7
15	ENFJ	5	8	9
16	ENTJ	4	7	8

From **Table 4**, it is concluded that testers, who have introvert personality characteristic are more appropriate for unit testing level as compared to extrovert personality characteristic, while the testers, who have extrovert personality characteristic are better in executing tasks related to integration and system testing level. Further it is also analyzed that testers with personality characteristics having J (Judging) dimension are better at executing tasks related to System testing level, means ENFJ, ESFJ, ESTJ, and ENTJ personality characteristic testers are suitable in executing system testing level than ENFP, ESFP, ESTP and ENTP personality characteristic testers, similarly testers with personality characteristic ISFJ, ISFJ, ISTJ and INTJ are suitable in executing system testing level than INFP, ISFP, ISTP and INTP personality characteristic testers.

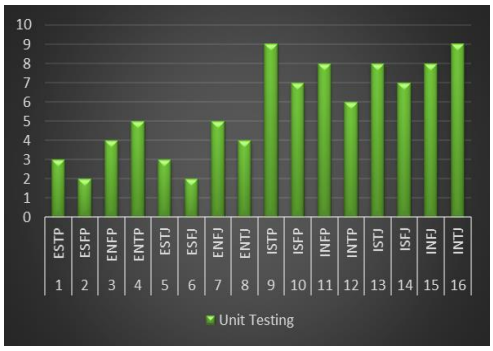


Fig. 2. Unit testing level Results by Personality Characteristics

Fig. 3. Integration testing level Results by Personality Characteristics

Fig. 4. System testing level Results by Personality Characteristics

From **Fig. 2**, it is shown that testers, who have introvert personality characteristic are appropriate for unit testing level as compared to extrovert personality characteristic, and from **Fig. 3** and **4**, it is shown that the testers, who have extrovert personality characteristic are better in executing tasks related to integration and system testing level.

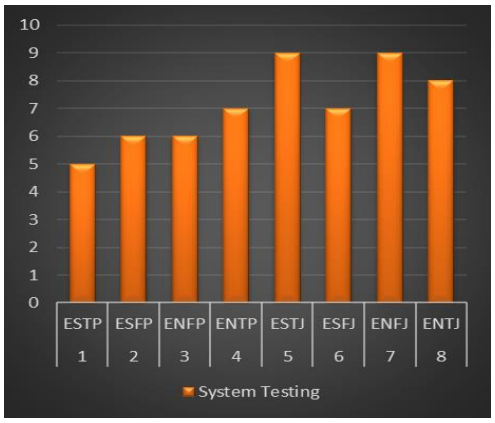


Fig. 5. System testing level Results by Extrovert Personality Characteristics

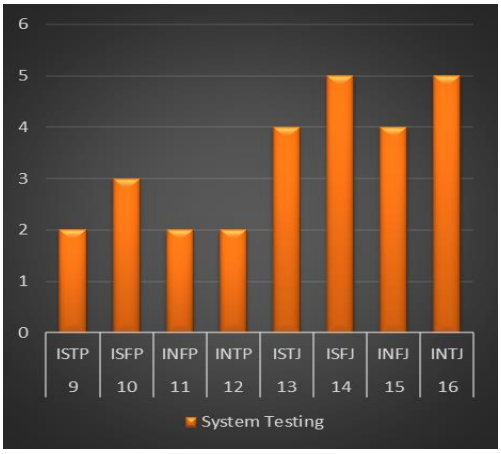


Fig. 6. System testing level Results by Introvert Personality Characteristics

From **Fig. 5**, it is shown that extrovert personality characteristics having J (judging) dimension executes System Testing Level improved than extrovert personality characteristics who do not have J (judging) dimension and from **Fig. 6**, it is shown that introvert personality characteristics having J (judging) dimension executes System Testing Level tasks improved than introvert personality characteristics who do not have J (judging) dimension.

5. Conclusion

The primary goal of this study was to govern the association between software testing level and tester with different personality characteristics, to achieve the results that shows which personality characteristic testers are appropriate at which software testing level. Software testing level related tasks should be allocated to related personality characteristic tester. From this research study it is concluded that testers who have introvert personality characteristic i.e., ISFP, ISTP, INFP, INTP, ISFJ, ISTJ, INTJ or INFJ are more appropriate for unit testing level as compared to extrovert personality characteristic, while the testers, who have extrovert personality characteristic i.e., ESFP, ESTP, ENFP, ENTP, ESFJ, ESTJ, ENTJ or ENFJ are better in performing tasks related to integration and system testing level. Further it is also analyzed that testers with personality characteristics having J (Judging) dimension are better at executing tasks related to System testing level, means ENFJ, ESFJ, ESTJ, and ENTJ personality characteristic testers are suitable in executing system testing level than ENFP, ESFP, ESTP and ENTP personality characteristic testers and similarly, testers with personality characteristic ISFJ, ISFJ, ISTJ and INTJ are suitable in executing system testing level than INFP, ISFP, ISTP and INTP personality characteristic testers.

References

[1] Ke Mao, Licia Capra, Mark Harman, and Yue Jia, “A Survey of the Use of Crowdsourcing in Software Engineering,” *in Journal of Systems and Software*, January 8, 2016.

[2] IEEE Guide for Software Verification and Validation Plans," *in IEEE Std 1059-1993* , vol., no., pp.1-87, 28 April 1994.

[3] Zainab Umair Kamangar, Umair Ayaz Kamangar, Qasim Ali, Isma Farah, Shahzad Nizamani and Tauha Husain Ali, “To enhance Effectiveness of Crowdsource Software Testing by applying Personality Types,” *in ICSIE '19 8th International Conference on Software and Information Engineering, Cairo, Egypt*, April, 2019.

[4] Yu, Dunhui & Zhou, Zhuang & Wang, Yi. "Crowdsourcing Software Task Assignment Method for Collaborative Development" *in IEEE Access*, PP. 1-1, March 2019.

[5] Leicht, N.; Rhyn M.; Hansbauer G., "Can laymen outperform experts? The effects of user expertise and task design in crowdsourced Software Testing," *in Twenty-Fourth European Conference on Information Systems (ECIS 2016),* İstanbul, Turkey, pp. 8, 2016.

[6] M. Allahbakhsh, B. Benatallah, A. Ignjatovic, H. R. Motahari-Nezhad, E. Bertino, and S. Dustdar, ''Quality control in crowdsourcing systems: Issues and directions,'' *in IEEE Internet Comput*., vol. 17, no. 2, pp. 76–81, Mar. 2013.

[7] G. Kazai, J. Kamps, and N. Milic-Frayling, "Worker types and personality traits in crowdsourcing relevance labels," *in Proc. CIKM, Scotland, U.K*, pp. 1941-1944*.,* Oct. 2011.

[8] L. F. Capretz, D. Varona, and A. Raza, "Influence of personality types in software tasks choices," *in Comput. Hum. Behavior,* vol. 52, pp. 373-378, Nov. 2015.

[9] Muhammad Zahid Tunio, Haiyong Luo, Cong Wang, Fang Zhao, Abdul Rehman Gilal, Wenhua Shao, "Task Assignment Model for Crowdsourcing Software Development: TAM," *in Journal of Information Processing Systems*, March, 2018.

[10] Muhammad Zahid Tunio, Haiyong Luo, Cong Wang, Fang Zhao, Abdul Rehman Gilal, Wenhua Shao, "Impact of Personality on Task Selection in Crowdsourcing Software Development: A Sorting Approach" *in IEEE Access*, Vol. 5, 2017.

[11] S. Cruz, F. Q. B. da Silva, and L. F. Capretz, "Forty years of research on personality in software engineering: A mapping study," *in Comput. Hum. Behavior*, vol. 46, pp. 94-113, May 2015.

Ensemble Learning Models for Classification and Selection of Web Services: A Review

[1]Muhammad Hasnain, [2]Imran Ghani and [3]Seung Ryul Jeong
[1]School of Information Technology, Monash University Malaysia,
[2]Department of Mathematics and Computer Sciences, Indiana University of Pennsylvania, Indiana, PA 15705, USA
[3]Graduate School of Business IT School of Management Information Systems Kookmin University, Seoul, Korea
[muhammad.malik1@monash.edu, ighani@iup.edu, srjeong@kookmin.ac.kr]
*Corresponding author: Muhammad Hasnain

Abstract

This paper presents a review of the ensemble learning models proposed for web services classification, selection, and composition. Web service is an evolutionary research area, and ensemble learning has become a hot spot to assess web services' earlier mentioned aspects. The proposed research aims to review the state of art approaches performed on the interesting web services area. The literature on the research topic is examined using the preferred reporting items for systematic reviews and meta-analyses (PRISMA) as a research method. The study reveals an increasing trend of using ensemble learning in the chosen papers within the last ten years. Naïve Bayes (NB) and 'Support Vector Machine' (SVM) and other classifiers were identified as widely explored in selected studies. Core analysis of web services classification suggests that web services' performance aspect can be investigated in future works. This paper also identified performance measuring metrics, including accuracy, precision, recall, and f-measure, widely used in the literature.

***Keywords*:** web services composition, quality improvement, class imbalance, machine learning.

1. Introduction

Ensemble learning combines multiple classification techniques to achieve a high prediction accuracy than the single classification technique. In other words, an ensemble learning model is represented by the collection of different classification techniques and aggregated results to derive a best-fit model [1]. An ensemble learning model is either the combination of the same class of classification techniques or different classification techniques. The former type of ensemble is a homogenous ensemble learning model [2], and the latter type is known as a heterogeneous ensemble learning model [3]. For instance, a boosting classification technique similar to a 'random forest' (RF) technique can be used to build a more accurate and robust classification model that involves multiple iterations. As AdaBoostM1, one of the boosting classifiers, undergoes more iterations, an ensemble learning model's predictive capacity is improved.

Ensemble learning models have been widely used in research areas, including bio-informatics [4], social media [5], software defect prediction, and classification [6]. Still, very little has been known about the applications of ensemble learning methods in web services' evolutionary research.

Web services selection regarding users' feedback helps determine users' trust. This trust of users is determined by evaluating the web services' quality attributes. The most well-known and widely explored quality attributes include response time, throughput, and reliability [7]. A user is always interested in selecting the best quality web service [8]. Quality attributes achieved at users' end need to be matched with the values mentioned in the 'service level agreement' (SLA) document [9]. The matching of users' achieved values and mentioned in the SLA document increases users' satisfaction. Different approaches have been proposed to evaluate web services classification and selection. To the best of our knowledge, this review is the first study that aims at reviewing the ensemble learning models and their applications in the area of web services.

The remainder of this study is as follows:

Section 2 describes the background; Section 3 describes the review method, section 4 presents results and discussion; and section 5 concludes the study and presents a few implications.

2. Background

The cost-sensitive and ensemble-based method (COSENS) framework [10] is the combination of ensemble methods with cost-sensitive learning. The proposed framework has been evaluated on the risk factors of outsourced software projects. A low misclassification cost and a high accuracy indicated that the COSENS framework established the new rigorous evaluation standard. For the COSENS framework, the selection of optimal classifiers was undertaken by using a non-cost-sensitive situation. Based on the selection criteria, SVM1 remained the most optimal classifier with 74.9% accuracy compared to NB and 'decision tree' (DT) classification techniques.

Huda et al. [11] proposed an oversampling ensemble method to address the class imbalance issue in predicting software defects. The researchers in this study focused on reducing the consequences of high 'false negative' (FN) instances. Defect prediction based on the over accuracies has lesser impacts on reducing the error rate.

The 'average probability ensemble' (APE) [12] incorporates several classification techniques (RB, 'gradient boosting' (GB), 'stochastic gradient descent' (SGD), NB, and SVM). The proposed ensemble model enhanced by the feature selection showed better performance in handling of redundant features and irrelevant features. The performance of base classification techniques was done by averaging individual classifiers' performance that reduced the error rate. The proposed APF technique was evaluated on six publicly available datasets that outperformed the existing approaches based on the single SVM and RF classifiers.

Ensemble learning [13] is proposed from bagging and boosting classification techniques. Boosting techniques perform better for text classification than the bagging techniques. Subsequently, the AdaBoost classification technique is chosen for ensemble learning. Since the AdaBoost technique enhances a weak classifier's performance, the capability to process data is also increased. In the latter technique, different weak classifiers are trained on the dataset and then combine them to achieve a robust classifier.

3. Review Methodology

To perform a rigorous review, preferred reporting items for systematic review and meta-analysis (PRISMA) guidelines [14] consisting of a 27-item checklist was used. The PRISMA flow chart is showing us the different phases of the review, such as identification, screening, eligibility, and Included (see **Fig. 1**). In the following, we present the search protocol adopted in each phase of the PRISMA guidelines to conduct the review on ensembles learning models and their applications in web services, classification, and selection research areas.

Table 1. Search protocol

SP	Data sources	ScienceDirect, IEEE Xplorer, Taylor and Francis, ACM Digital Library, and Springer.
	Search string	"Ensemble learning models" OR "Ensemble classification models" AND "Web services classification OR Web services selection."
	Search fields:	Title, abstract, keywords, and full contents.
	Period	From January 1, 2010, to August 31, 2020.
	Documents	Journal and Conference paper
	Language	English

Table 1 is the illustration of the search protocol (SP) adopted in this study. As listed in **Table 1**, we selected five popular digital libraries as data sources that covered machine learning and web services topics. Second, we constructed the search string based on the objectives of this study. Third, we applied eligibility criteria: (1) Search fields included title, abstract, and keywords; (2) Data coverage of studies ranged between January 1, 2010, and August 31, 2020; (3) Documents were Journal and conference papers; and (4) Language that considered studies published in English.

3.1 Research Question

Based on the reviewed literature, we aim to investigate the state of the art approaches at the ensemble learning models and web services. We proposed to define a few research questions as follows:

RQ1-What are state of the art approaches that employed ensemble learning models in the context of web services classification and selection?

RQ2- What are the challenges of the ensemble learning models in the studied approaches?

RQ3- What are the performance metrics reported in the chosen studies?

3.2 Inclusion and Exclusion Criteria

Inclusion criteria

- We include papers that discuss ELM along with web services themes.
- We include papers written only in the English language.

Exclusion Criteria

- We exclude papers with only abstracts.
- We exclude papers that discuss topics other than web services classification, selection, composition, and discovery.
- We exclude duplicate papers.

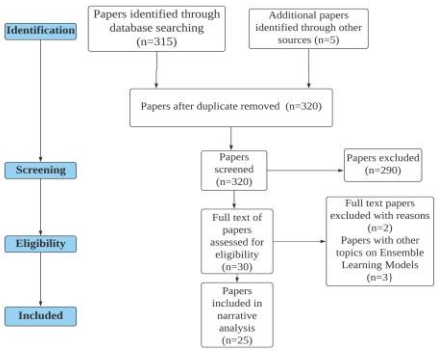


Fig. 1. Review strategy

We report the four steps review strategy in Fig. 1. In the first step, we show the number of studies identified from search strings. In the next step, we perform the screening of papers and exclude those papers, which do not follow the studies' inclusion criteria. The subsequent step is focused on assessing the documents with the availability of full text. Only papers that accept the inclusion criteria are accepted.

4. Results and Discussion

This section presents answering the proposed RQs and results in the discussion.

RQ1-What are state of the art approaches that employed ensemble learning models in the context of web services classification and selection?

To answer RQ1, we have investigated the chosen studies.

We find several studies that answer RQ1. A study [15] in recent research proposed to use Bayesian Information Criteria (BIC) along with Akaike Information Criteria (AIC) as an evaluation metric for prediction of response time and throughput values. The latter mentioned approach used invocation time of web services with similar functionalities and then forecasted the best web services with convincing quality of services (QoS) values. Before this research, [16] proposed the weighted voting ensemble of eleven classifiers to select Amazon web services and Microsoft services, providers. Earlier works did not present the criteria to know how users are satisfied with the services. Users' feedback information of web services was further used to evaluate the performance of ensemble classification models. The logistic regression model showed better performance than the ensemble learning models.

Web services discovery [17] is another real challenge for researchers and software practitioners to develop service-oriented web applications. This is further compounded by the limited number of services management frameworks and intelligent categorization. To capture the functional semantics, researchers proposed a machine learning technique to overcome the earlier mentioned limitations. Ensemble learning classification categorizes the services into specific domains. AdaBoost, along with Random Forest and Bagging classifiers,

were used. Combined techniques helped to capture the semantic relatedness between services. Accuracy results of ensemble learning models were better than the based estimators such as SVM and NB classifiers.

The proposed approach for the classification of web services uses the significant vote based ensemble classification models, including Naïve Bayes, SVM, and decision tree (J48) [18]. On the other hand, the proposed approach in [24] combined Naïve Bayes, Random Forest, and SVM classification techniques. The former technique used a decision tree classifier, and the latter used the Random Forest classifier. The rest of the two classifiers, such as Naïve Bayes and SVM, were the same for two ELM techniques. The ensemble learning technique resulted in higher accuracy in comparison with the single technique. Higher accuracy results from ELM were also based on the improved preprocessing. An earlier approach [24] employed the Naïve Bayes, SVM, and REPTree classifiers in an ensemble learning fashion and showed accuracy improvement in the calibration of 'web services description language' (WSDL) documents. However, classification accuracy in [18, 24] was better than the study [25]. This might be due to the maturity of ensemble learning techniques in the current era of machine learning.

Missing value and imbalance data of attributes is another challenge in the bioinformatics domain. To overcome this issue, the study [20] proposed a service-oriented support decision system (SOSDS) based on the services oriented architecture. Web services represent the model for problem construction and related functionalities. Decision making is performed by using lousy quality data and then assuming it to a binary classification problem. A committee of the SVMs is treated as an oracle that aims to correct training data and enable rule-based learning. Recent research [21] determined the prediction of unknown QoS values by building the K-nearest neighborhood of services based on their similarity results. So unknown values in the QoS matrix were generated using the known ones of neighbors of the concerning services. A nonnegative latent factor (NLF) model performed better in predicting the missing QoS values with high accuracy.

Anti-patterns are known as bad practicing solutions from the evolution or modification in web services with poor QoS values [22]. Earlier detection of anti-patterns in web services is performed by the ensemble learning method. The 'best training ensemble' (BTE) approach performed better than the 'majority voting ensemble' (MVE).

Web services composition plays a significant role in satisfying the users' preferences [23]. Best composition patterns' mining with less time is still the central research area. The ensemble of the 'best first decision tree' (BFDT) along with extreme boosting models is proposed in recent research. The proposed approach is aimed to extract and mine the users' interesting patterns. Classification of web patterns into positive and negative correlations helps correct users' most interesting patterns.

RQ2- What are the challenges of the ensemble learning models in the studied approaches?

A study [15] shows a better prediction of the QoS metric values but does not show the efficiency compared with the recent becoming popular approaches. The subjective and manual labeling of data proposed is time-consuming [19]. The subjective data labeling and its accuracy depend on the quality of labeling data. In [16], we noticed that subjective labeling had time consumption and quality labeling issues. Therefore, we suggest researchers develop a precise method for labeling data that can reduce time in labeling large datasets.

The proposed approach [17] is efficient for the classification of web services with a static dataset. This approach shows limitations in managing the newly added services to a central repository known as 'universal description, discovery, and integration' (UDDI). Thus, automatic management may be activated to handle the changes occurring in web services.

ELM proposed in [20] can help solve the class imbalance or missing values problem, but still, we need to address noisy labels.

Classification accuracies values reported in [24] are highly impressive in classification aspects. However, the application of the proposed ELM technique can be used to predict web services' performance. Based on the performance results, software practitioners may know the future performance of their web services and correct if performance lowers than the promised QoS values.

We can conclude that ensemble learning models still have the challenges of manual data handling, i.e., labeling, noisy labels, time consumption.

RQ3- What are the performance metrics reported in the chosen studies?

To answer RQ3, we have extensively studied the chosen papers. We have listed widely explored performance measuring metrics in the following **Table 2**.

Table 2. Performance measuring metric

Sr.#	Performance metrics	Reference
1.	Sensitvity	[16][20]
2.	Specificity	[16][20]
3.	Cohen's Kappa	[16][24]
4.	Matthews' Correlation Coefficient	[16]
5.	Recall Precision and F-Measure	[17][22][24]
6.	Accuracy	[17][20][24] [25]
7.	Mean Absolute Error (MAE)	[20][21]
8.	Root Mean Square Error (RMSE)	[20][21]
9.	Area under the ROC Curve (AUC)	[22]
10.	Web Pattern Identification Accuracy (WPIA)	[23]

Table 2 shows a list of ten performance measuring metrics studied in the chosen papers. The accuracy-based metrics such as accuracy, precision, recall, and f-measure were examined in studies [17, 20, 22, 24, 25] followed by the rest of the metrics. Only two papers [20. 21] studied MAE and RMSE metric. Other than these metrics, Mathews' correlation coefficient, AUC, and WPIA metrics were used in [16], [22], and [23] studies, respectively.

5. Conclusions

This paper presents an overview of the literature on the research topic of ensemble learning models and web services classification, selection, and composition. Advancement in machine learning technology introduces new challenges to web services' research area due to the imbalanced classes and missing values, decreasing the accuracy of classification models. This paper shows an increasing trend in the ELM concerning web services quality improvement. In total, this paper identified ten performance measuring metrics. We grouped precision, recall, and f-measure because the majority of studies report all three metrics.

The findings of this study have implications for researchers, web services developers, and managers. The results suggest that automated data labeling can save time and cost for larger web systems classification.

References

[1] K. W. Walker and Z. Jiang, "Application of adaptive boosting (AdaBoost) in demand-driven acquisition (DDA) prediction: A machine-learning approach," *The Journal of Academic Librarianship*, vol. 45, no. 3, pp. 203-212, 2019.

[2] M. Hosni, A. Idri, and A. Abran, "Evaluating filter fuzzy analogy homogenous ensembles for software development effort estimation," *Journal of Software: Evolution and Process*, vol. 31, no. 2, p. e2117, 2019.

[3] M. Tang, X. Dai, J. Liu, and J. Chen, "Towards a trust evaluation middleware for cloud service selection," *Future Generation Computer Systems*, vol. 74, pp. 302-312, 2017.

[4] A. Anaissi, M. Goyal, D. R. Catchpoole, A. Braytee, and P. J. Kennedy, "Ensemble feature learning of genomic data using support vector machine," *PloS one*, vol. 11, no. 6, p. e0157330, 2016.

[5] O. Araque, I. Corcuera-Platas, J. F. Sánchez-Rada, and C. A. Iglesias, "Enhancing deep learning sentiment analysis with ensemble techniques in social applications," *Expert Systems with Applications*, vol. 77, pp. 236-246, 2017.

[6] Ö. F. Arar and K. Ayan, "Software defect prediction using cost-sensitive neural network," *Applied Soft Computing*, vol. 33, pp. 263-277, 2015.

[7] J. Liu, M. Tang, Z. Zheng, X. Liu, and S. Lyu, "Location-aware and personalized collaborative filtering for web service recommendation," *IEEE Transactions on Services Computing*, vol. 9, no. 5, pp. 686-699, 2015.

[8] H. A. Kadhim and H. N. Nawaf, "Improve the Accuracy of Dirichlet Reputation System for Web Services," in *2018 11th International Conference on Developments in eSystems Engineering (DeSE)*, 2018: IEEE, pp. 78-82.

[9] J. Heinermann and O. Kramer, "Machine learning ensembles for wind power prediction," *Renewable Energy*, vol. 89, pp. 671-679, 2016.

[10] Y. Hu *et al.*, "Cost-sensitive and ensemble-based prediction model for outsourced software project risk prediction," *Decision Support Systems*, vol. 72, pp. 11-23, 2015.

[11] S. Huda *et al.*, "An ensemble oversampling model for class imbalance problem in software defect prediction," *IEEE access*, vol. 6, pp. 24184-24195, 2018.

[12] I. H. Laradji, M. Alshayeb, and L. Ghouti, "Software defect prediction using ensemble learning on selected features," *Information and Software Technology*, vol. 58, pp. 388-402, 2015.

[13] L. Yuan-jie and C. Jian, "Web service classification based on automatic semantic annotation and ensemble learning," in *2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum*, 2012: IEEE, pp. 2274-2279.

[14] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement," *Int J Surg*, vol. 8, no. 5, pp. 336-341, 2010.

[15] V. P. Singh, M. K. Pandey, P. S. Singh, and S. Karthikeyan, "An Econometric Time Series Forecasting Framework for Web Services Recommendation," *Procedia Computer Science*, vol. 167, pp. 1615-1625, 2020.

[16] M. Raza, F. K. Hussain, O. K. Hussain, M. Zhao, and Z. ur Rehman, "A comparative analysis of machine learning models for quality pillar assessment of SaaS services by multi-class text classification of users' reviews," *Future Generation Computer Systems*, vol. 101, pp. 341-371, 2019.

[17] S. S. Kamath and V. Ananthanarayana, "Semantics-based web service classification using morphological analysis and ensemble learning techniques," *International Journal of Data Science and Analytics*, vol. 2, no. 1-2, pp. 61-74, 2016.

[18] U. Qamar, R. Niza, S. Bashir, and F. H. Khan, "A majority vote based classifier ensemble for web service classification," *Business & Information Systems Engineering*, vol. 58, no. 4, pp. 249-259, 2016.

[19] M. S. Graham, I. Drobnjak, and H. Zhang, "A supervised learning approach for diffusion MRI quality control with minimal training data," *NeuroImage*, vol. 178, pp. 668-676, 2018.

[20] M. Zięba, "Service-oriented medical system for supporting decisions with missing and imbalanced data," *IEEE journal of biomedical and health informatics*, vol. 18, no. 5, pp. 1533-1540, 2014.

[21] X. Luo, M. Zhou, Y. Xia, Q. Zhu, A. C. Ammari, and A. Alabdulwahab, "Generating highly accurate predictions for missing QoS data via aggregating nonnegative latent factor models," *IEEE transactions on neural networks and learning systems*, vol. 27, no. 3, pp. 524-537, 2015.

[22] S. Tummalapalli, L. Kumar, and N. B. Murthy, "Prediction of Web Service Anti-patterns Using Aggregate Software Metrics and Machine Learning Techniques," in *Proceedings of the 13th Innovations in Software Engineering Conference on Formerly known as India Software Engineering Conference*, 2020, pp. 1-11.

[23] D. G. Chakravarthy and S. Kannimuthu, "Extreme Gradient Boost Classification Based Interesting User Patterns Discovery for Web Service Composition," *Mobile Networks and Applications*, vol. 24, no. 6, pp. 1883-1895, 2019.

[24] N. Negi, P. Chaudhary, and S. Chandra, "Web service classification based on non-functional parameters using vote based classifier," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 23, no. 1, pp. 45-55, 2020.

[25] L. Yuan-jie and C. Jian, "Web service classification based on automatic semantic annotation and ensemble learning," in *2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum*, 2012: IEEE, pp. 2274-2279.

KSII The 12th International Conference on Internet (ICONI) 2020.

An Automatic Approach to Identifying Tasks from a Code Change

Taeyoung Kim and Suntae Kim
Department of Software Engineering, Chonbuk National University
Seoul, South Korea
[e-mail: rlaxodud1200@jbnu.ac.kr, stkim@jbnu.ac.kr]
*Corresponding author: Suntae Kim

Abstract

A software maintainer generally consults the code changes between the current and the previous version of the code and examines which part of the code has been modified in the code repository. However, generally there are multiple tasks such as adding a new feature or refactoring in a code change, so that the maintainer spend a lot of time trying to find the related code changes. In order to handle this issue, this paper suggests an automatic approach to identifying tasks from a code change. It first starts with generating AST (Abstract Syntax Tree) parsing tree pair of previous and current versions for all classes existing in the two commits. Then, the change type of each AST element is analyzed by comparing two parsing trees. At the last step, the tasks are identified by the suggested algorithm. For the evaluation, we applied our approach into our project and identified 54 tasks from randomly selected 10 commits. Then, we analyzed the appropriateness of the identified tasks and obtained 0.81 of precision and 0.77 of recall respectively.

***Keywords*:** Software maintainer, Code Change, AST

1. Introduction

A software maintainer spends a lot of time analyzing legacy source code in order to fix bugs or add new features [1]. During the analysis, the maintainer generally consults the code changes between the current and the previous version of the code and examines which part of the code has been modified, added or deleted in the commit of the code repository. Based on the code change analysis, the maintainer tries to understand the software developer's intention for the code changes [2]. However, most of developers tend to update the code for performing diverse tasks (e.g., adding or deleting features, and fixing bugs) at the same time and upload the all code changes in a single commit into a code repository [3]. This single commit containing diverse modified resources for multiple tasks makes a software maintainer difficult to understand a code change.

There have been two major research streams for facilitating understanding of the code changes. One stream is to record the key typing and replaying them later (see, [4], [5], [6]). This approach records the code typing activities in the IDE, and replays them later like a video clip. Although they give detailed information to the developer, the vast amount of those information makes a maintainer feel tedious and hard to understand the developer's intention of the code changes. Another stream handles the code changes in terms of code change elements such as classes and methods, and visualizes code change elements based on the AST parsing tree (see, [7], [8]). However, these approaches are only achievable when the system is executable with the all associated libraries in the development environment.

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT)(No. 2020R1F1A1072039).

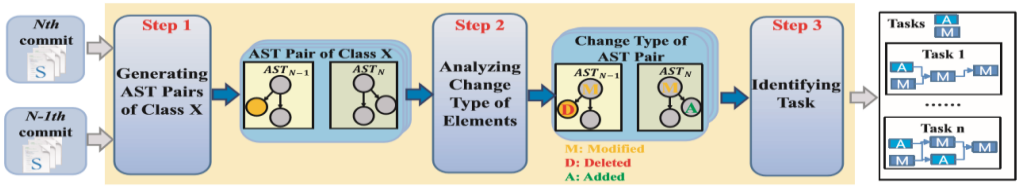


Fig. 1. An Overview for Identifying Tasks from a Code Change

In order to address the issue, we propose a novel approach to automatically identify tasks from the code changes only with two commits from the code repository. The task is the cause of a code change composing of adding or removing the related code elements such as classes, methods, attributes and statements. The examples of a task are adding a new feature and refactoring. Identifying tasks first starts with generating AST parsing tree pair of previous and current versions of all classes existing in two commits. Then, the change type (e.g., *Added* or *Modified*) of each AST element is analyzed by comparing two parsing trees. At the last step, candidate tasks are identified and merged into a complete task by our suggested algorithm. For the evaluation, we applied our approach into our Java based project *MISRA-C Rule Checker* and identified 54 tasks from randomly selected 10 commits. Then, we analyzed the appropriateness of the identified tasks and obtained 0.81 of precision and 0.77 of recall respectively.

The remainder of the paper is organized follows: Section 2 introduces the definition of the task and steps for identifying tasks. Section 3 shows identified tasks and quantitative evaluation of our approach. We concludes this paper and discuss future work in Section 4.

2. Identifying Tasks from a Code Change

Fig. 1 shows an overview of our approach to identifying tasks from a code change. Inputs of our approach are *Nth* and *N-1th* commits from the code repository. Based on the pair of the commit, it first starts with generating AST parsing tree pair of the two versions for all classes in the two commits. At the second step, the change type of each AST element is analyzed by comparing two parsing trees. From the change types, candidate tasks are identified and merged them into a complete task by our suggested algorithm at the last step. From the next subsection, after defining the task and its elements, detailed explanation on each step is described.

2.1 Definition of Task

As mentioned earlier, a task is the cause of the code changes, and all activities for constructing our system such as implementing new features, modifying or deleting existing features, fixing bugs, and refactoring can be representative examples. A task is a minimal element for unit testing and code sharing (i.e., commit to code repository). The task is completed by consecutively adding or removing the related code elements such as *Class* and *Statement* because a change of one code element causes the subsequent changes of other elements. Thus, a task is composed of the set of added/modified/deleted code elements and their causal relationship of the code change in a time order. However, the missing information of these chronological and causal relationships in the code repository makes a maintainer difficult to understand the code change.

Table 1. TYPE OF TASK ELEMENTS

Task element	Description	Notation
Class	Class, Test Case Class	C, TC
Attribute	Attribute	A
Method	Method, Test Case Method	M, TCm
Statement	Statement, Test Case Statement	Stmt, TCstmt

Expression	Description	Notation
Method Invocation	Invoke owned or other classes methods	C, TC
Class instance creation	Instantiates a class by invoking the constructor	A
Field access	Access owned attributes	M, TCm

Before going further, we first define the elements of a task as shown in **Table 1**, composing of four elements Class, Method, Attribute, and Statement, each of which is very general in the object-oriented programming paradigm. The reason that the test case should be specially cared is that a test case is a good starting point to identify tasks because passing a test

case is inevitable to complete a task. The statement indicates the control statement such as if and for, containing diverse expressions including *Conditional Expression*, Array Creation and etc. Among the expressions, we selected three major expressions *Method Invocation*, *Class instance creation*, *Field Access*, each of which is used as a good clue to identify tasks. In addition, the following relationships exist among task elements:

- $\mathrm{C_i} = \{\mathrm{set}(\mathrm{A_{i,j}}), \mathrm{set}(\mathrm{M_{i,j}})\}$
- $\mathrm{TC_i} = \{set(\mathrm{TCM_{i,j}})\}$
- $\mathrm{M_{i,j}} = \{\mathrm{set}(\mathrm{Stmt_{i,j,k}})\}$
- $\mathrm{Tcm_{i,j}} = \{\mathrm{set}(\mathrm{TCstmt_{i,j,k}})\}$
- $\mathrm{Stmt_{i,j,k}}, \mathrm{TCstmt_{i,j,k}} = \{set(\mathrm{MI_i}), set(\mathrm{CI_i}), set(\mathrm{FA_i})\}$
- $\mathrm{CR_{X,Y}} = \{X, Y \in \{C, TC, A, M, TCm, Stmt, TCstmt\}\}$

In the relation, the set indicates a group of the designated element. Also, CR is Causal Relationship between two aforementioned elements, denoting change of the first element X causes other element Y 's changes. For example, CR(*foo(),bar()*) implies that *add/delete/modify* of the method *foo()* have an influence on modification of the method *bar()*. Based on the definition, the tasks are identified from the code changes, which is composed of task elements and casual relationship.

2.2 Step 1. Generating AST Pairs

The purpose of this step is to generate AST parsing tree pair of the two versions for all classes in two commits. The input of this step is a pair of resources from the *Nth* and *N-1th* commits of the code repository. As each commit contains a set of classes, comparing two sets of resources from two commits is necessary to identify the set of added, modified, and deleted task elements. Two inputs can be fetched from the code repository[1]. Then, we generate AST parsing tree of all classes from the *Nth* and *N-1th* commits, and make AST Pairs of the same class in the two sets from the two commits[2]. In the case of a new or deleted class of the *Nth* commit, there is only one part in the AST Pair.

[1]We fetched code commits from Github and used the JGit [9] library to fetch the resources of the specific commit.

[2] Eclipse JDT AST [10] is used to build the AST parsing tree.

2.3 Step 2. Analyzing Change Types of Elements

Each task element has one change type among *Added*, *Modified*, and *Deleted*. While the *Added* type denotes that a task element is newly added in the *Nth* commit which does not exist in the N-1th commit, the *Deleted* type indicates that the task element is deleted in the *Nth* commit, so that it does not exist in *Nth* commit. The *Modified* type is analyzed by comparing AST Pair of a class existing in the two commits. Although the change type of a class is *Modified*, the it's owned attributes, methods, and statements of the class can be other change types like other task elements.

When analyzing the change type of a method, the method signature is extracted by using the regular expression and compared the methods that has the same signature from two commits. When the method signature is same but the body part is different, its change type is defined as *Modified*. For an attribute and a statement, we first tokenize them into words. For an attribute, its name as well as accessibility (e.g., public/private) and its data type are also tokenized. For the statement, the only condition part of the statement are targeted for the tokenization, because the body part of the statement is handled by another statement.

Based on the two sets of words of an attribute or a statement from the *Nth* and *N-1th*, the similarity is computed by Equ.1, where TE indicates the task elements and *N* and *N-1* denote the commit numbers. When the similarity is less then 30%, we set Modified into the change type of the task element[3]. When it exceeds 30%, we set Added into the task elements of the *Nth* commit and Deleted into that of the *N-1th* commit. This second step results in all task elements annotated its change type.

- $\mathrm{Sim}(TE_N, TE_{N-1}) = \frac{set(Word)_{TE_N} \cap set(Word)_{TE_{N-1}}}{set(Word)_{TE_N} \cup set(Word)_{TE_{N-1}}}$

2.3 Step 3. Identifying Tasks

Given the AST Pairs and its change types of the task elements from the second step, this step is to identify tasks by using the suggested algorithm. Alg. 1 shows the suggested approach to identifying candidate tasks composing of task elements and change type annotations. The input parameter of the algorithm is the AST Pair of a specific commit, so that the algorithm should be

[3] We obtained 30% as a change type threshold by sensitive analysis.

called twice. One is called with AST Pair of the *Nth* commit, and another is called with that of N-1th commit. The output of the algorithm is a set of candidate tasks of a *Nth* and *N-1th* commits. Then, two outputs are merged based on the task element name.

Alg. 1 starts with initializing all sets of task elements (see line 3). Then, it is divided into task identification starting from test cases (see line 4-10) and from normal classes (see line 12-21) in the AST Pair. Task identification from test cases starts first, which means that the task elements related to a test case tend to be clearly identified as a task if the test case is modified or newly added. The algorithm executes *Search&Build* algorithm as shown in Alg. 2 (see line 4-8). In order to build the causal relationship between two task elements, the *Search&Build* algorithm is suggested. In the algorithm, the expression sets (*set(MI), set(CI) and set(FA)*) are extracted from the statement and assigned into the temporal *set(X)*. Then, each expression set is passed into the *BuildCR* algorithm to perform iteratively building the causal relation between elements of statements and other sets *set(M)* and *set(A)* as shown in line 2-4.

Algorithm 1: Identifying Tasks

```
1  IdentifyTasks (ASTPair):
2      Initialize Tasks = {};
3      Initialize set(C), set(M), set(A), set(TC),
         set(MI), set(CI), set(FA);
4      foreach TC_i ∈ set(TC) do
5          foreach TCm_{i,j} ∈ set(TCm_i) do
6              foreach TCstmt_{i,j,k} ∈ set(TCstmt_{i,j})
                 do
7                  Search&Build(TCstmt_{i,j,k},
8                  set(M), set(A), Tasks);
9              end
10         end
11     end
12     foreach C_i ∈ set(C) do
13         foreach M_{i,j} ∈ set(M_i) do
14             if M_{i,j} is not used then
15                 foreach Stmt_{i,j,k} ∈ set(Stmt_{i,j})
                     do
16                     Search&Build(Stmt_{i,j,k},
17                     set(M), set(A), Tasks);
18                 end
19             end
20         end
21     end
22     return Tasks;
```

Algorithm 2: Search & Connect Causal Relationships

```
1  Search&Build
     (Stmt_{i,j,k}, set(M), set(A), Tasks):
2      Initialize
         set(X) ← {set(MI), set(CI), set(FA)};
3      foreach X_l ∈ set(X) do
4          BuildCR(set(X_l), set(M), set(A), Tasks);
5      end
```

Alg. 3 finds all causal relationships and identify candidate tasks. It should be noted that *set(X)* in this algorithm is a group of expressions of the statement from the *Search&Build* algorithm. Each expression in *set(X)* such as method invocation and class instance creation are checked if they have a causal relationship through *CheckCR()* function. Then, if it meets the criteria regarding the causal relationship, it establishes the causal relationship between the statement and method or attribute (see line 4-6 and 15-16). Also, as the method contains more statements in the body, it starts the *Search&Build* algorithm with the method body statement again, which is denoted in the line 7-10. Note that once method and attributes are found as a task element, their class are correspondingly found because relation between the two elements has been already analyzed through the AST parsing tree.

When building the causal relationship in the *CheckCR()* function, the criteria according to the expression is used as shown in Table 2. For the method invocation, if the method identifier of the method invocation expression is the same with that of the *set(M)* and their parameter size of the method declaration is same and their change types are Modified, they are considered as the two elements have a causal relationship.

Algorithm 3: Build Causal Relationships

```
1  BuildCR set(X_l), set(M), set(A), Tasks):
2      foreach X_{l,m} ∈ set(X_l) do
3          foreach M_{i,j} ∈ set(M) do
4              if CheckCR(X_{l,m}, M_{i,j}) then
5                  CR ← (Stmt_{i,j,k}, M_{i,j});
6                  Tasks ← CR(Stmt_{i,j,k}, M_{i,j});
7                  foreach
                     MStmt_{i,j,k} ∈ Set(MStmt_{i,j}) do
8                      Search&Build(MStmt_{i,j,k},
9                      set(M), set(A), Tasks);
10                 end
11             end
12         end
13         foreach A_{i,j} ∈ set(A) do
14             if CheckCR(X_{l,m}, A_{i,j}) then
15                 CR ← (Stmt_{i,j,k}, A_{i,j});
16                 Tasks ← CR(Stmt_{i,j,k}, A_{i,j});
17             end
18         end
19     end
```

It means that the change of the task element (i.e., a method declaration in this case) have an influence on the change of the method invocation expression. Similar to this, the causal relationships of the class instance creation and field access expression also can be identified. Note that Alg. 1 is called twice with AST Pair of the Nth and N-1th commit. Each call produces two sets of the

candidate tasks and they are merged if they have the same task elements.

Table 2. CRITERIA FOR A CAUSAL RELATION IN CheckCR()

Expression	Criteria #1	Criteria #2
MI(Method invocation)	Method Identifier	Parameter size
CI(Class instance creation)	Method Identifier == Class Identifier	Parameter size
FA(Field access)	Attribute Accessor == Attribute Identifier In a Class	

3. Evaluation

This section presents an evaluation of our approach. First we show the result of tasks automatically identified by our suggested approach. Then, we measure the quality of the identified task by computing precision and recall. For this evaluation, we selected our project *MISRA-C Rule Checker* opened in the Github [4] because background knowledge regarding the domain and source code is necessary to validate the identified task for this initial phase research. *MISRA-C Rule Checker* is our open source project that supports verification of the automotive software source code according to the MISRA-C 2004 coding standard [11]. In the project, five developers were involved and made 250 commits in the code repository during five months. From this project, we randomly selected 10 commits and tried to identify tasks.

3.1 Tasks Identified by Our Approach

From the 10 commits, we have identified 54 tasks. Due to the space limit, this paper only shows four identified tasks from the #c90d312 commit[5] as shown in **Fig. 2**. In the figure, The change type Added, Modified, Deleted are expressed in the green, yellow and red oval, and causal relationships between the task elements are expressed as arrow line in the task. Task #1 is composed of three modified task elements in a MisraPreferencePage, and the three task elements are connected with two causal relationships. In the class, the data type of the attribute tableViewer has been modified from TableViewer to CheckboxTablewViewer. We can understand that the change of the attribute causes changes of the addTableListener and createTable methods. Thanks to the time order display, a maintainer can easily understand the order of the developers work.

Similar to this, Task #2 is composed of one added task element, the contributeButtons() method, and one modified task element the createContents() method. It implies that the createContents() method has been modified because of the new contribution button. Task #3 and #4 are identified as independent tasks as they were modified without any causal relationship with the elements, which is fixing bugs that does not trigger any change of another part of the code or refactoring of functions included in the existing test case.

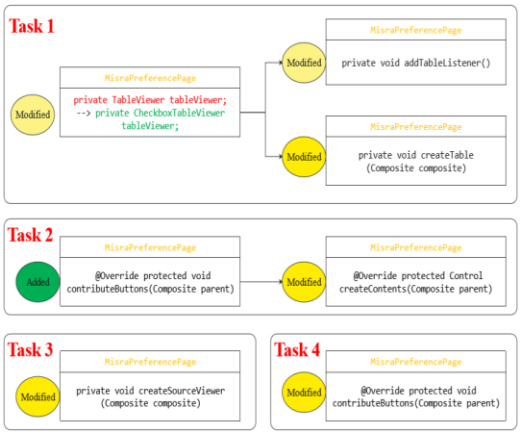


Fig. 2. the Tasks of Commit c90d312

3.2 Analysis of Identified Tasks

For the quality of the identified tasks, we have quantitatively analyzed the results of our approach. For the analysis, five developers involved in the project manually identified tasks from the *diff* of the Github and checked the validity of the identified tasks and the causal relationships between task elements. Then, we obtained the precision and recall of our approach and summarized that in Table 3. Our approach identified 54 tasks. Among the 54 tasks, the developers confirmed that 44 tasks are appropriate and 10 tasks are not. Thus, we obtained 0.81 of precision on average. In addition, the developers found 57 tasks, which is three more tasks rather than our approach. Thus, we obtained 0.77 of recall. For the validity of the causal relationships, we obtained the 0.85 of precision and 0.89 of recall respectively.

Most of the reasons for the invalid task identification were because of our callee search

[4] MISRA-C Rule Checker Homepage available at https://github.com/stkim123/kr.ac.jbnu.ssel.misrac

[5] https://github.com/stkim123/kr.ac.jbnu.ssel.misrac/commit/c90d312b20 646255c85ba9f006dcb69443046291

approach of the method invocation that just find out a callee method matched with the same method identifier and the same number of parameters. When there is a class hierarchy, our approach to finds the callee of the method invocation found the only super class method, though there are many changes in the methods of the child classes. Those changes of the child classes are identified as wrong tasks. This is caused by the none-type binding feature of the static analysis, which implies that we can only recognize the exact data type at runtime. This reason is the same reason of the wrong identification of the causal relationships. Another reason is that two change types of the task elements are *Modified*, but one modification are not caused by another modification.

Table 3. THE PERFORMANCE OF IDENTIFYING TASKS AND CAUSAL RELATIONSHIP

Commit	Type	Precision	Recall
e5a962f	Task	0.67()	0.50
	Causal Relationship	0.77	0.83
85d17d	Task	1.00	1.00
	Causal Relationship	1.00	1.00
d8f7593	Task	0.33	0.50
	Causal Relationship	0.67	0.67
1e44398	Task	1.00	1.00
	Causal Relationship	1.00	0.50
c624ae4	Task	0.67	0.50
	Causal Relationship	0.77	1.00
c90d312	Task	1.00	1.00
	Causal Relationship	1.00	1.00
e5a5285	Task	1.00	1.00
	Causal Relationship	1.00	1.00
7ebd4e6	Task	1.00	1.00
	Causal Relationship	1.00	1.00
2b6513	Task	1.00	1.00
	Causal Relationship	1.00	1.00
64a2fbf	Task	0.54	0.46
	Causal Relationship	0.66	0.66
Average	Task	0.81	0.77
	Causal Relationship	0.85	0.89

4. Conclusions

This paper presented an automatic approach to identifying tasks from a code change. In order to achieve the purpose, we proposed the definition of the task and three steps to identify the tasks: 1) generating AST Pairs of a class, 2) analyzing a change type of the task element, and 3) identifying tasks with the proposed algorithm. For the evaluation of our approach, we compared the tasks identified by our approach and the developers involved in our project, and obtained 0.81 of precision and 0.77 of recall. For the future work, we have a plan to improve our algorithm for identifying tasks and apply them into open source projects to show feasibility of our approach.

References

[1] Tobias Baum, Kurt Schneider, and Alberto Bacchelli. 2017. On the Optimal Order of Reading Source Code Changes for Review. In Proceedings of the 2017 IEEE International Conference on Software Maintenance and Evolution. 329–340.

[2] A. Von Mayrhauser and A. M. Vans, "Program comprehension during software maintenance and evolution," Computer, vol. 28, no. 8, pp. 44–55, 1995.

[3] T. Mens. 2002. A State-of-the-Art Survey on Software Merging. IEEE Transactions on Software Engineering 28, 5 (2002), 449–462.

[4] R. Robbes and M. Lanza, "Spyware: A change-aware development toolset," in Proceedings of the 30th international conference on Software engineering, 2008, pp. 847–850.

[5] L. Hattori and M. Lanza, "Syde: a tool for collaborative software development," in Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering-Volume 2, 2010, pp. 235–238.

[6] J. Park, Y. H. Park, S. Kim, and A. Oh, "Eliph: Effective visualization of code history for peer assessment in programming education," in Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, 2017, pp. 458–467.

[7] K. Maruyama, E. Kitsu, T. Omori, and S. Hayashi, "Slicing and replaying code change history," in Proceedings of the 27th IEEE/ACM International Conference on Automated Software Engineering, 2012, pp. 246–249.

[8] T. Kamiya, "Code difference visualization by a call tree," in 2018 IEEE 12th International Workshop on Software Clones (IWSC). IEEE,

2018, pp. 60–63.
[9] The JGIT website. [Online]. Available: https://git-scm.com/book/de/v2/Appendix-B\%3A-Git-in-Ihre-Anwendungen-einbetten-JGit
[10] The JDT website. [Online]. Available: https://www.eclipse.org/jdt/core/r2.0/dom\%20ast/ast.html
[11] T. Doering. Misra-c:2004. [Online]. Available: http://caxapa.ru/thumbs/468328/misra-c-2004.pdf

Business Process Feature Model in Identifying Potential Core Assets

Ain Balqis Md Nor Basmmi, Shahliza Abd Halim and Dayang N. A. Jawawi
School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia
Skudai, Johor, Malaysia
[e-mail: ainbalqis96@gmail.com, shahliza@utm.my, dayang@utm.my]
*Corresponding author: Ain Balqis Md Nor Basmmi

Abstract

It has been recognized that the combination of Software Product Line (SPL) and Service Oriented Architecture (SOA) can facilitate variability modelling of services in a systematic manner. An explicit variability modelling in complex organizations tends to offer advantages in structuring activities and processes effectively including during core assets development for system reuse. Business Process Feature Model (BPFM) has been introduced in existing works to aid in developing business process in product line and is believed to encompass all business activities thus making it the most expressive business process model for system reuse. However, the current approach of core assets development does not properly tackle the concerns in systematic reuse through BPFM. Therefore, this paper proposes the use of BPFM in identifying only potential reusable core assets that contributes in solving the problem of maintaining the structural stability of a system as more services are being reused. BPFM is generated using a tool and is assigned a weight according to its probability of occurrence in potential configurations. The weighted BPFM will later be used in identifying Trade-off Analysis between the structural stability of system and priority of services depending on the probability of occurrence to derive a set of potential reusable core assets. A case study is represented in this research to demonstrate the technique. From this study, a set of potential reusable core assets is specified by utilizing Trade-off Analysis through BPFM. Trade-off Analysis through BPFM has proven to be feasible for identifying potential reusable core assets systematically.

***Keywords*:** Business Process Feature Model, Software Product Line, Service Oriented Architecture, Trade-off Analysis

1. Introduction

The advent and development of technology has had a huge impact on companies in different sectors of the industry. The environment of business is marked by a high level of change and most companies must adapt to these changes in order to retain competitiveness which acquire solution as quickly as possible [1]. However, delivering the need of each company requires high efforts in adapting the demand of requirements and traditional method of building system from scratch surely induced time and cost. Therefore, a systematic reuse approach during core assets development must be adopted to achieve a flexible, quick, and cost-effective system development. Service Oriented Product Line (SOPL) has appeared as a synergy between Service Oriented Architecture (SOA) and Software Product Line (SPL) to aid a systematic reuse approach that can lead to product flexibility, faster market, and maintenance saving [2].

With SOPL, system is developed by reusing

This research was supported by a research grant from UTM Encouragement Grant (UTMER)

existing assets in a systematic manner. As the main focus is to find a technique that can develop a system with similar services without structuring it from scratch, a modelling notation that can support variability is needed. Business Process Feature Model (BPFM) was introduced as a notation that combines concept coming from both feature modelling and business process modelling in a new notation [3] and is believed to encompass all business activities thus making it the most expressive business process model for system reuse. Some researchers have already uncovered BPFM as an approach to deal with variability of business process but not to the extent of applying BPFM as a notation to aid in core assets development during system reuse [4, 5]. Therefore, the main contribution of our ongoing research is to derive a set of potential reusable core assets by using BPFM.

In this study, BPFM is generated by using a modelling environment from BPFM webpage [6] and is weighted according to its probability of occurrence in potential configurations [7]. As more services are being reused, to maintain the structural stability of a system is a challenge that need to be catered. Thus, Trade-off Analysis between structural stability of system and priority of services depending on the probability of occurrence is a way to derive only a set of potential core assets to be reused systematically and at the same time to maintain the stability of a system to be developed.

The rest of this paper is structured as follows: Section 2 examines related work. Section 3 describes the proposed approach based on the enhancement from existing work. Section 4 presents the implementation of proposed approach in detail. Last but not least, Section 5 presents the lesson learned and concludes this paper with some recommendations for future research.

2. Related Works

As an emerging approach, SOPL has drawn interest of the research community. However, core assets development approach through SOPL that uses BPFM to represent activities of business process variant is not adequately addressed. Most existing studies uses BPFM to express variability of business process [3, 4] but not to the extent of deriving potential reusable core assets. Therefore, this study focuses on applying Trade-off Analysis using BPFM. Research by [8] provides a promising reference in terms of method on identifying potential core assets with Trade-off Analysis between stability and value but the implementation of stability in their study is too simplistic which does not concern in other factors such as dependencies between services. An addition of services is a factor that contributes in the decrease of structural stability of the system [8]. Especially when the added services consist of services with high probability of occurrence and high priority. Thus, our works differs in the aspect that we focus on providing a technique to derive reusable assets that is justified by considering its priority and probability of occurrence through Trade-off Analysis.

3. Identifying Potential Reusable Core Assets

Proposed technique in this study is designed by enhancing the existing work [8] with addition of BPFM that act as a UML activity diagram to model process families [9]. The BPFM is modelled using graphical editor via BPFM tool from the functionality by ADOxx development platform [10]. ADOxx is a set of tools developed to facilitate the implementation of meta-model-based modelling environments [3, 11, 12].

Then, a weight is assigned to each member of the BPFM based on the probability of occurrence to form a Weighted Business Activity Set (WBAS) that was proposed by [7]. WBAS consist of several variability type of business activities such as Mandatory and Optional that are outside of variant region, Variant with cardinality, Mandatory Variant, and Optional Variant. Each of it carries different weight as shown in Fig. 1. WBAS is a way to aid the derivation of potential reusable core assets during Trade-off Analysis later.

Variability type of business activity	Weight	Description
Mandatory (outside variant region)	1	
Optional (outside variant region)	$\frac{1}{2}$	
Variant of a simple variation point	$\frac{\sum_{i=a}^{i=b}\binom{v}{i}}{2^v}$	With cardinality $(a..b)$ and total "v" variants
Mandatory variant in a variant region	$\frac{1}{2}$	Region as a whole may be optionally selected
Optional variant in a variant region	$\frac{1}{4}$	Same as above

Fig. 1. Probability of Occurrence for WBAS [7]

In terms of Trade-off Analysis, [8] provides the most promising reference in deriving potential reusable core assets through SOPL. Trade-off Analysis in [8] focuses on finding the best compromise between added value and the structural stability of a system. The value is allowed to be defined in different ways such as a charged price for a service or its priority. Since it is challenging to get a market price for a particular service, this study uses a priority of services that was defined by depending on its probability of occurrence or WBAS, and preferred services by user. High probability of occurrence indicates a high priority of service and it should be noted that different type of service differs in level of priority and ease of reuse. Hence, this study addresses the question of how to determine which services to be reused that could potentially be core assets by taking into account the constraint of maximizing the priority of service and minimizing a decrease in structural stability of a system. **Fig. 2** depicts process of the proposed technique.

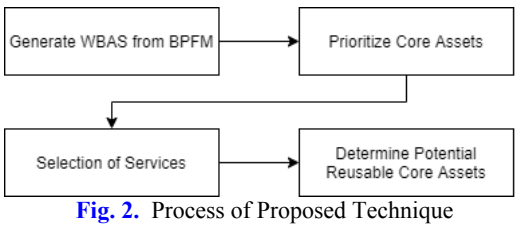


Fig. 2. Process of Proposed Technique

4. Implementation and Result

The proposed systematic reuse technique to identify a set of potential reusable core assets is outlined accordingly in this section.

4.1 Case Study

The case study illustrated in this study is based on Patient Navigation Program (PNP) from [13]. PNP is an intervention that shows promise to improve colorectal cancer screening adherence and reduce disparities among cancer patients [14]. In Malaysia, PNP functioned as data management, analysis, barrier prediction, milestone prediction, incidence prediction, efficacy, response, and side effects of the treatment analysis and manual reporting [15]. However, there are only one PNP that has been implemented in Malaysia's hospital which is in Hospital Tengku Ampuan Rahimah (HTAR). To ensure that patients receive the most effective and systematic treatment across Malaysia, the use of PNP need to be widen. The focus of the case study in this paper is only on services offered by the PNP as a proof-of-concept for the proposed technique in this study. Some of the services offered includes: (1) File Tracking, (2) Status, (3) Notification, (4) Milestones, (5) Assessment, and (6) Case Data.

4.2 Generate WBAS from BPFM

All services are identified using the feature model of PNP and is then transformed into BPFM (as shown in **Fig. 3**) by using ADOxx development tool and is assign a weight to form a WBAS as shown in Table 1. BPFM is used as the main technique of service identification as it combines business activities that is needed to be performed and the functionality of a system. Table 1 summarizes all services that will be used to represent this case study and identified weight for each service.

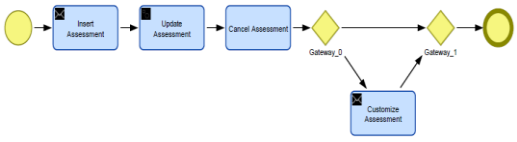


Fig. 3. BPFM for Assessment Activity

Table. 1. List of Services Offerred

Service ID	Services	WBAS Value
File Tracking		
s1	Update Location	1
s2	Movement History	1
Status		
s3	Update	1
Notification		
s4	Insert	1
s5	Update	1
Milestone		
s6	Preset	1
s7	Insert	1
s8	Update	1

Assessment		
s9	Customize	0.5
s10	Insert	1
s11	Update	1
s12	Cancel	1
Case Data		
s13	CRUD	1
s14	Clinical Diagnosis	1
s15	Treatment Plan	1
s16	Visit Details	1
s17	Pre-Surgery Info	1
s18	Post-Surgery Info	1
s19	Other Assessment	1
s20	Treatment	1
s21	Neo-Adjuvant Therapy	0.47
s22	Adjuvant Therapy	0.47
s23	Radiotherapy	0.47
s24	2nd Chemo	0.5
s25	3rd Chemo	0.25
s26	Hormone Therapy	0.47
s27	Targeted Therapy	0.47
s28	Side Effects	0.5
s29	Oncology Referral	1

4.3 Prioritize Core Assets

For the purpose of this study and as a proof-of-concept, three users with their respective preferred services is included as depicted in **Table 2**. Ticks indicate that a services preferred by user shall be included in the composition of service instances.

Table. 2. List of Services Preferred

Service ID	S1 (User 1)	S2 (User 2)	S3 (User 3)
s1	/		/
s2		/	
s3	/		/
s4	/		
s5		/	/
s6			/
s7	/		
s8		/	
s9	/		/
s10		/	
s11			/
s12	/		
s13	/	/	/
s14	/		
s15	/		/
s16		/	/
s17	/		
s18		/	/
s19			
s20	/	/	/
s21	/		
s22		/	
s23		/	/
s24		/	
s25			/
s26			/
s27	/		
s28		/	
s29			/

Core assets is prioritized according to value of WBAS. Higher WBAS indicates a high priority. As there are multiple services with same weight, the order is randomly chosen. Noted that during prioritization of core assets, only services that was not chosen by users in **Table 2** is picked. It specifies services to be added aside from services that are most preferred by users. **Table 3** outlines the services ranked in descending order for each instances.

Table. 3. Rank of Services

Rank	S'1	S'2	S'3
1	s2	s1	s2
2	s5	s3	s4
3	s6	s4	s7
4	s8	s6	s8
5	s10	s7	s10
6	s11	s11	s12
7	s16	s12	s14
8	s18	s14	s17
9	s19	s15	s19
10	s29	s17	s24
11	s24	s19	s28
12	s28	s29	s21
13	s22	s9	s22
14	s23	s21	s27
15	s26	s26	-
16	s25	s27	-
17	-	s25	-

4.4 Selection of Services

In this step, we established the ultimate balance between rising a composition's value and lowering stability. We claim that a decline in stability must be justified by added services that was prioritized. Structural stability must be maintained in order to avoid increment of cost and effort in delivery and maintenance of services. Hence, to identify the best compromise between structural stability of a system and prioritized added services is needed. This results to the use of Trade-off Analysis. Normalization of AAV dan Stability is carried out and as an example, the results for Instance 1 is shown in **Table 4**.

Table. 4. Trade-off Analysis for Instance 1

i	Added service	AAV Increased (normalized)	Stability Decreased (normalized)
1	s2	norm(1) = 0	norm(0.93) = 1
2	s2,s5	norm(2) = 0.09	norm(0.87) = 0.87
3	s2,s5,s6	norm(3) = 0.17	norm(0.81) = 0.76
4	s2,s5,s6,s8	norm(4) = 0.26	norm(0.76) = 0.66
5	s2,s5,s6,s8,s10	norm(5) = 0.34	norm(0.72) = 0.57

When adding s2, the stability loss is estimated as 1-[1/(13+1)] = 0.93. For the addition of s2, s1, and s4, the loss of stability is 1-[3/(13+3)] = 0.81. The values are then normalized. Value for instance 1 and 2 are calculated in a similar manner and the result of Trade-off Analysis for each instance is illustrated in **Fig. 4**, **Fig. 5**, and **Fig. 6** respectively.

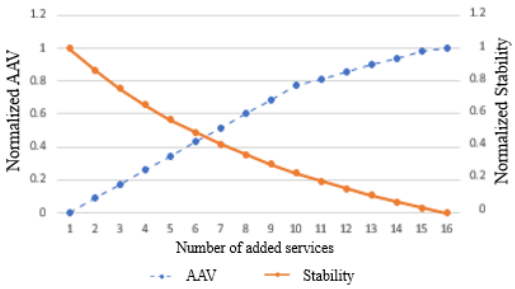


Fig. 4. Trade-off Analysis Instance 1

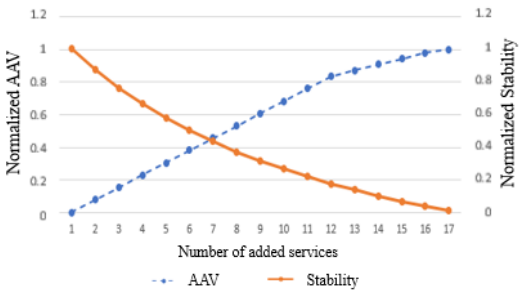


Fig. 5. Trade-off Analysis Instance 2

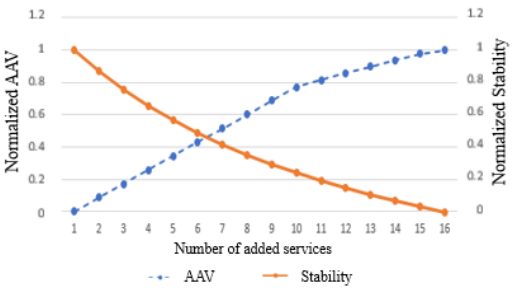


Fig. 6. Trade-off Analysis Instance 3

4.5 Determine Potential Reusable Core Assets

In choosing a set of potentially reusable core assets, only the best trade-off analysis result is selected which are those with a close value between AAV and Stability. Therefore, only the first seven results are being selected for each instance. S”1, S”2, and S”3 after adding services are as shown in **Fig. 6**.

S”1 = S1 U {s2,s5,s6,s8,s10,s11,s16}
= {s1,s3,s4,s7,s9,s12,s13,s14,s15,s17,s20,s21,s27}
U {s2,s5,s6,s8,s10,s11,s16}
= **{s1,s2,s3,s4,s5,s6,s7,s8**,s9.s**10,s11,s12,s13**,s14,
s15,**s16**,s17,**s20**,s21,s27}

S”2 = S2 U {s1,s3,s4,s6,s7,s11,s12}
= {s2,s5,s8,s10,s13,s16,s18,s20,s22,s23,s24,s28}
U {s1,s3,s4,s6,s7,s11,s12}
= **{s1,s2,s3,s4,s5,s6,s7,s8,s10,s11,s12,s13, s16**,
s18,**s20**,s22,s23,s24,s28}

S”3 = S3 U {s2,s4,s7,s8,s10,s12,s14}
= {s1,s3,s5,s6,s9,s11,s13,s15,s16,s18,s20,s23,s25,
s26,s29} U {s2,s4,s7,s8,s10,s12,s14}
= **{s1,s2,s3,s4,s5,s6,s7,s8**,s9,s**10,s11,s12,s13**,s14,
s15,**s16**, s18,**s20**,s23,s25,s26,s29}

Fig. 6. Result of Potential Reusable Core Assets

As a conclusion, a set of potential reusbale core assets, Sca = S”1 n S”2 n S”3. The services which overlap in these instances are considered as candidates for core assets of SOPL. These are the services that existed or could be in existing instances in all current instances S"1, S"2, and S"3 without requiring multiple changes and leading to an increase in value at the same time. It guarantees that {s1, s2, s3, s4, s5, s6, s7, s8, s10, s11, s12, s13, s16, s20}. will effectively gotten to be instances of a recently created product line.

4. Conclusions

In this paper, we presents a Trade-off Analysis technique between the added services based on its priority and structural system stability. From this study, it is possible to systematically reuse services provided by the PNP. Trade-off Analysis through BPFM is feasible for selecting only the potential set of core assets to be reused systematically. It can be assured that priority of services and structural stability of a system are two measures that should not be missed as more services added might contributes in a loss of stability. Therefore, the result implies that only

services with high priority should be reused. As future work, a comprehensive empirical evaluation on the reusability of identified core asset will be our main focus.

Acknowledgement

This material is based upon work supported by the UTM-TDR Grant Vot No. 06G23.

References

[1] Ezenwoke, A., Misra, S., and Adigun, M. O., "An approach for e-commerce on-demand service-oriented product line development," *Acta Polytechnica Hungarica*, Vol 10, No. 2, pp. 69-87, 2013.

[2] Castelluccia, D. and Boffoli, N., "Service-oriented product lines: a systematic mapping study," *ACM SIGSOFT Software Engineering Notes*, Vol 39, No. 2, pp. 1-6, 2014.

[3] Cognini, R., Corradini, F., Polini, A. and Re, B., "Extending feature models to express variability in business process models," in *International Conference on Advanced Information Systems Engineering*, pp. 245-256, 2015.

[4] Cognini, R., Corradini, F., Polini, A. and Re, B., "Business process feature model: an approach to deal with variability of business processes," in *Domain-Specific Conceptual Modeling*, pp. 171-194, 2016.

[5] Azouzi, S., Brahmi, Z., and Ghannouchi, S., "Customization of multi-tenant learning process as a service with Business Process Feature Model," *Procedia Computer Science*, Vol 126, pp. 606-615, 2018.

[6] Business Process Feature Model, Retrieved from https://austria.omilab.org/psm/content/bpFM/download?view=download.

[7] Khoshnevis, S. and Shams, F., "Automating identification of services and their variability for product lines using NSGA-II," *Frontiers of Computer Science*, Vol 11, No. 3, pp. 444-464, 2017.

[8] Galster, M. and Eberlein, A., "Identifying potential core assets in service-based systems to support the transition to service-oriented product lines," *18th IEEE International Conference and Workshops on Engineering of Computer-Based Systems*, pp. 179-186, 2011.

[9] Diego, D., "Integrating PPI Variability in the Context of Customizable Processes by Extending the Business Process Feature Model," in *EDOC*, 2020.

[10] ADOxx Metamodelling Platform, Retrieved from https://www.adoxx.org.

[11] Fill, H. G., and Karagiannis, D., "On the conceptualisation of modelling methods using the ADOxx meta modelling platform," *Enterprise Modelling and Information Systems Architectures (EMISAJ)*, Vol 8, No. 1, pp. 4-25, 2013.

[12] Kuhn, H., "The ADOxx Metamodelling Platform," in *Workshop on Methods as Plug-Ins for Meta-Modelling*, 2010.

[13] Ushananthiny, S., "Reusable Patient Navigation Software Architecture using Software Product Line Approach," UTM, Malaysia, 2018.

[14] DeGroff, A., Coa, K., Morrissey, K. G., Rohan, E., and Slotman, B "Key considerations in designing a patient navigation program for colorectal cancer screening," *Health promotion practice*, Vol 15, No. 4, pp. 483-495, 2014.

[15] Suvelayutnan, U., "Reusable Patient Navigation Software Architecture using Software Product Line Approach," in *Tractates Sabbat and 'Eruvin*, 2018.

Extended Feature Model for Quality Attribute Identification in Software Product Line Architecture

'Afifah M. Ashari[1], Shahliza A. Halim[1], Dayang N. A. Jawawi[1], and Ushananthiny Suvelayutnan[2]
[1] School of Computing, Faculty of Engineering
Universiti Teknologi Malaysia
[2] Hospital Sultanah Aminah, Jalan Persiaran Abu Bakar Sultan,
80100 Johor Bahru, Johor – Malaysia
afifahashari95@graduate.utm.my, dayang@utm.my, shahliza@utm.my, unanthiny0405@gmail.com
*Corresponding author: 'Afifah M. Ashari

Abstract

Software Product Line (SPL) is one of the means that help with the rapid growth of system developments. SPL is an approach that promotes planned reusability where the product line domain is well analyzed, and the core assets derived strategically for the predicted reuse. Nevertheless, the lack of quality attribute (QA) representation becomes a problem that causes difficulties and time for QA's identification process to be increased. In SPL engineering, features express the variabilities and commonalities among programs. Therefore, the representation of QA in feature model can help improve the process of identification. This research focuses on studying how QA's presentation in the early phase of software development in SPL work. In this paper, QA is being implemented in feature model of SPL development to achieve the study objective. The extended feature model will be applied in Patient Navigation Program domain to help in the wide spread of the system.

***Keywords*:** Extended Feature Model, Software Product Line

1. Introduction

There is an importance on taking account of all the requirements, including quality attribute requirements while working on domain engineering phase as it determines the customer satisfaction towards the product. Unfortunately, unclear description of quality attribute (QA) from the stakeholder makes the process harder to complete. To solve this, some researchers used feature model (FM) extension to represent the QA in early phase of software development. Noted the lack in representation of QA using FM with clear requirement description and interdependencies between QA with functional feature. Therefore, implementation of the extended feature model will be done in this chapter for quality identification on the domain of breast cancer PNP system.

In the early stages of software engineering, collecting user requirements is one of the most critical phases because the system needs to be produced with correct functions that can only be known by acquiring the necessary requirements. There are significant differences between non-functional requirements (NFR) and quality requirement. NFR describes the qualitative and quantitative measures of how the system must behave, while quality requirement is NFR that discuss product availability, maintenance,

This research was supported by Universiti Teknologi Malaysia – UTMHR Grant Vot No. 08G67

modifiability, timeliness, throughput, and responsiveness, reliability, security, and scalability [1]. Therefore, from the NFR collected from the user, the respective QA can be identified for the product development. In order to assist on the more complex identification of the QA, an extended feature model is proposed as representation for the QA in a FM.

This study aims to propose an extendable feature model with quality attribute representation. This paper is presented as follows: Background of the study is described in Section 2. The findings of the study are presented in Section 3 and the conclusions in Section 4.

2. Background

2.1 Feature Model

Features express the variabilities and commonalities among programs in a SPL. A feature model defines the valid combinations of features, where each combination corresponds to a program in an SPL. A feature model defines the correct combinations of features in a domain [2]. It is organized hierarchically and is graphically depicted as a feature diagram.

In order for systematic reusability to be done with SPL, feature model assist a lot in determine the features that needed to be used. There are other used of feature model in which it can be modify to adapt with the project or research. In the case of SPL architecture, it can be note that some researchers modified their feature model to cater to the problem of their studies which will be investigate in next section.

2.2 QA Representation in Feature Model

Several researchers are exploring the representation of QA in FM in the past as it is significant for architecture selection and in product line engineering. There is a noticeable lack of QA representation in architecture selection. The researchers solved a few issues regarding the identification and selection of QA for some products. In past paper [3], the researchers studied the QA assessment for configured products that they mentioned were neglected in most existing product configuration approaches. They also stated that there is a high-cost factor in fixing any problem that may arise if the produced product cannot meet the customers' requirements on quality attributes. Therefore, a target product's quality attributes should be assessed in the earliest stage of product development in software product lines—product configuration in a feature model [3].

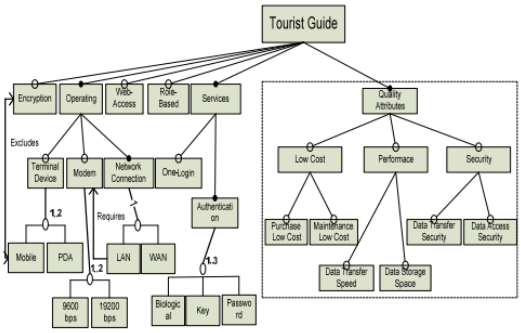


Fig. 1. Extension of Feature Model with QA Sub-Feature Tree by Zhang et al., 2011.

Looking at how the researchers constructed their EFM, to cater to the problem of difficulties and time-consuming activity of identifying QA, they decided to represent the QA in the FM by extending their current FM with a sub-feature tree to represent QA as shown in **Fig. 1**.

Other than that, other study [4] highlighted the need to elicit and refine QA requirements because QA has different and imprecise meanings depending on the domain. The researchers extend their feature model with the quality feature, consisting of a mechanism from Architecture Trade-off Analysis Method (ATAM) evaluations. It represents quality attributes, variability, and the influence on the quality of the functional architecture and implementation features. **Fig. 2** shows the extended feature model of the study.

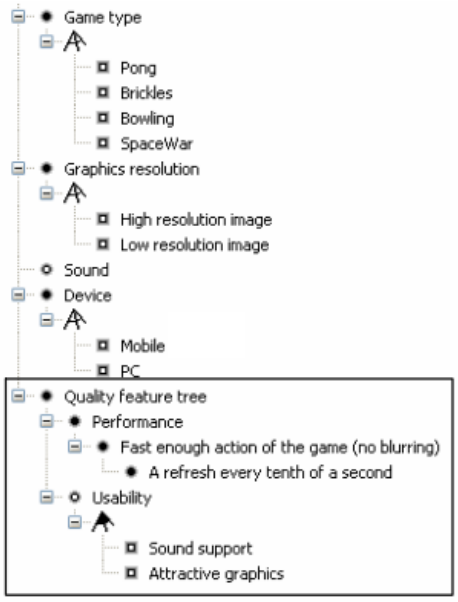


Fig. 2. Extension of quality feature in FM by Etxeberria et al., 2008

Both studies implemented an extended feature model to cater to the problem of QA representation for a product line. Nonetheless, a few improvements can be made to increase QA representation in feature model. Factors such as whether the QA is in the latest ISO standard, following the domain quality, and domain expert have to be considered for the extended feature model. Thus, an extended feature model specifically for PNP is required for this research proposed. **Table 1** summarize the advantage and disadvantage of the propose extended feature model of the two studies.

Table. 1. Advantage and Disadvantage of past Extended Feature Model

Ref.	Advantage	Disadvantage
[3]	Extend feature model with QA and functional feature	did not specify QA based on current ISO standard or domain of HIS
[4]	Non-Functional Requirement statements represent in Feature Model	

3. Extended Feature Model

In the previous section, the background of QA and FM have been described. In this section, the implementation of the FM as QA representation is being done. In this study, the main objective is to propose an extendable feature model with quality attribute representation. **Fig. 2** shows the summary of proposed method of EFM for PNP.

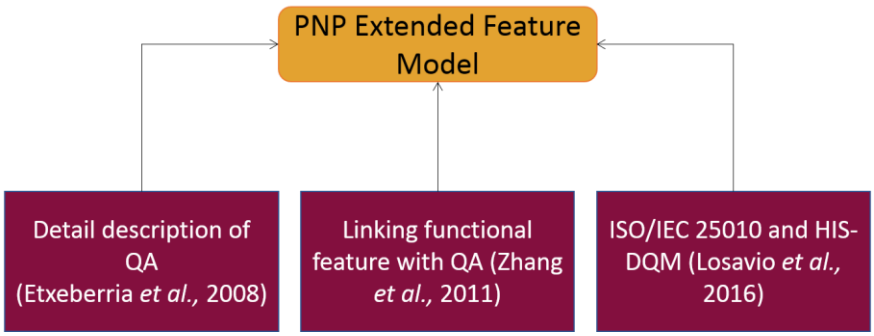


Fig. 3. Extension of Feature Model

According to the past section, two propose method on the use of Extended Feature Model (EFM) had been studied [3] [4]. Upon considering the advantage and disadvantage of both method, enhancement on the feature model had been done in this chapter. The enhancement is, the QA to extend on the feature model are mapped with the standard ISO which is ISO/IEC 25010 and domain quality model which is HIS-DQM. This is to ensure that the QA is within the PNP domain [5].

3.1 Implementation

The PNP FM is build using Feature IDE tools in Eclipse software system. Feature IDE tools in Eclipse provide a graphical editor where the EFM of PNP can be edited. Representation of QA in the FM will be implemented by extending PNP feature model with QA. An example of the FM is shown in Fig 4. Besides that, QA obtained from domain experts also being implemented to a sub-tree feature model. The QA sub-tree that will be added for the extension is shown in **Fig. 5**.

The next process is to enhance the EFM by mapping QA from the stakeholder with current standard ISO which is ISO/IEC 25010 and domain quality model which is HIS-DQM. This is to ensure that the QA is within the PNP domain. **Fig. 6** below shows the EFM of PNP with QA feature tree.

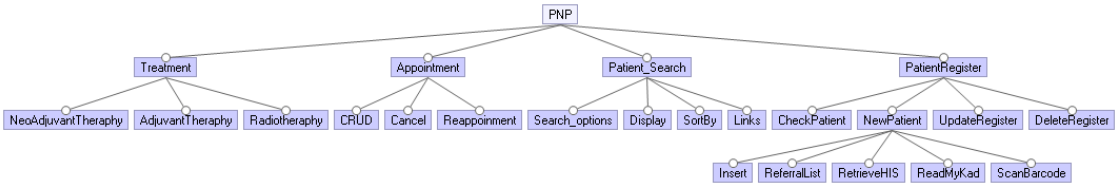


Fig. 4. Feature Model of PNP

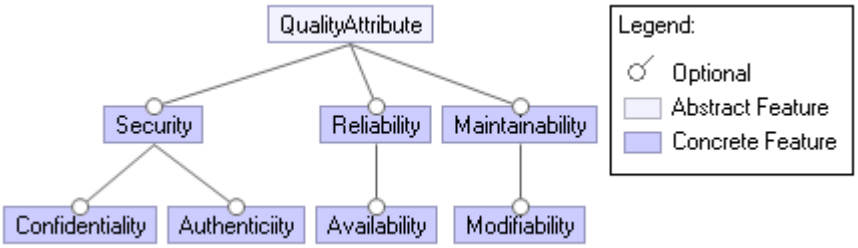


Fig. 5. Sub-tree of Quality Attribute

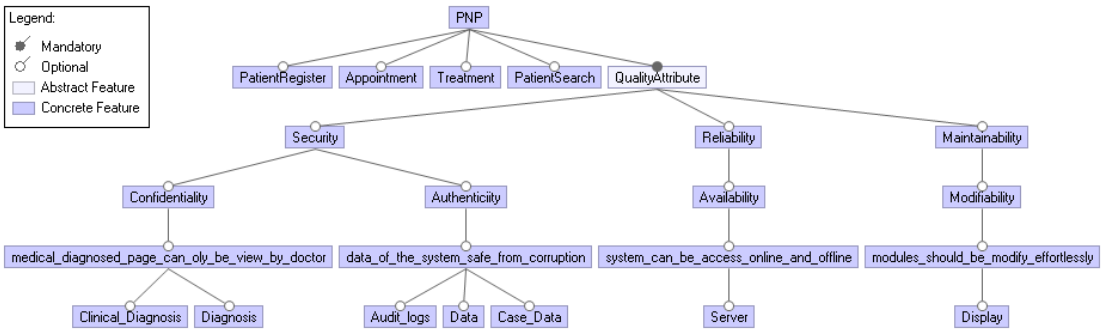


Fig. 6. Feature Model of PNP with QA extension

To check the validity of the method, an interview had been done with the stakeholder. Several questions had been asked towards the stakeholder to gather the respond on the research topics which are listed as below:

1. Are the Extended Feature Model easy to understand?
2. Are the Extended Feature Model easy to use?
3. Are the Extended Feature Model help in identifying NFR and QA feature for the system?

From the interview, several comments from the stakeholder are being noted such as that the extended feature model really do assist them to determine the non-functional requirements and quality attribute of the product. Clear description of the NFR also help domain expert to give relationship between the QA and functional features of the system. It is also stated that the extended feature model is easy to understand and easy to use. **Table 2** listed the detailed answer from domain expert who is the IT department staff of Hospital Tengku Ampuan Rahimah.

Table. 2. Domain Experts Validation

Validation Question	**Domain Expert Responds**
Are the Extended Feature Model easy to understand?	Yes it is easy to understand because of the clear and detail description of the feature model.
Are the Extended Feature Model easy to use?	Yes it is easy to used.
Are the Extended Feature Model help in identifying NFR and QA feature for the system?	It really help on identifying NFR and QA feature because of the detailed elicitation.

In conclusion, the responds from the domain expert was positive in term of the understandability of use because of the clear and detail description of the used of the feature model in the study. Other than that, it also assisted in identifying QA for the software product line architecture selection.

4. Conclusions

In this study, the EFM with QA representation PNP is implemented to assist in identifying QA in the early phase of SPL engineering. The EFM contained the QA with a detailed description to help reduce the unclear description of the QA. Besides that, the functional features for the QA were also displayed in the EFM to assist in the interdependencies of the system's components with the features. Further work needs to be done on testing the EFM by implementing it in other domains. Other than that, further evaluation to investigate the configuration of the core assets in the SPL architecture should be conducted to increase the efficiency of software development.

5. Acknowledgement

The authors wish to acknowledge Universiti Teknologi Malaysia for grant UTM-TDR Grant Vot No. 06G23, and Ministry of Higher Education (MOHE) for FRGS Grant Vot No. 5F117, which have made this research possible. The authors would also like to express their warmest appreciation on continuous support and inputs provided by members of Embedded Real Time Software Engineering Laboratory

References

[1] Nurbojatmiko, Eko K. Budiardjo, and Wahyu C. Wibowo. "Slr on Identification & Classification of Non-Functional Requirements Attributes, and Its Representation in Functional Requirements." *Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence - CSAI '18*, 2018.

[2] Thum, Thomas, Christian Kastner, Sebastian Erdweg, and Norbert Siegmund. "Abstract Features in Feature Modeling." *2011 15th International Software Product Line Conference*, 2011.

[3] Zhang, G., Ye, H., & Lin, Y. “Modelling Quality Attributes in Feature Models in Software Product Line Engineering”. *ICSOFT*, 2011.

[4] Etxeberria, Leire, and Goiuria Sagardui. "Variability Driven Quality Evaluation in Software Product Lines." 2008 *12th International Software Product Line Conference*, 2008.

[5] Losavio, Francisca, Oscar Ordaz, and Henry Márquez. "Assessment for quality product derivation from a software product line reference architecture." *Revista Antioqueña de las Ciencias Computacionales 5.2*, 2015.

Hybrid in-memory storage for stable and high-performance virtual desktop Infrastructure

Dae Won Kim, Soo Cheol Oh, Ji Hyeok Choi and Seong Woon Kim
Electronics and Telecommunications Research Institute. 218 Gajeong-ro, Yuseong-gu, Daejeon, South Korea
[e-mail: won22@etri.re.kr, ponylife@etri.re.kr, jihyeok_choi@etri.re.kr ksw@etri.re.kr]
*Corresponding author: Dae Won Kim

Abstract

In general, the virtual desktop infrastructure builds and manage a virtual machine in a cloud environment to access it remotely in order to use the user's desktop existing OS such as window, linux, mac, etc. Virtual machines for virtual desktops is created as virtual machine images on disk storage. In this case, the usage of high-speed storage (SSD, NVMe, etc.) is increasing due to the limitation of I/O performance, and there are even cases where a virtual machine image is maintained and used in main memory. Operating a virtual machine image on main memory can increase its performance compared to any other disk. However, RAM used as main memory is expensive and disappears after the system is shut down due to its volatile characteristic. Therefore, additional work is required to operate the virtual desktop on the main memory. In this paper, a hybrid in-memory storage is proposed to combine main memory and other disk such as a high-speed SSD to safely operate a virtual desktop on RAM (main memory) from the volatile characteristic. Also, the proposed storage shows 6 time faster for writing and 42 time faster for reading operation comparing with general disk.

***Keywords*:** cloud computing, VDI, virtual machine, In-memory disk, RAM, virtual desktop

1. Introduction

In a virtual desktop system [1], virtual disks can be stored in main memory to increase the performance of virtual machines [2]. However, this method has a disadvantage in that this method cannot store a lot of data because the capacity of the main memory is limited. When several virtual machines are created and operated on a single physical machine at the same time, the load is concentrated on the disk storage where the virtual disk is located and the performance is degraded.

In addition, when the virtual machine image is allocated as a user's disk in a virtual desktop, the size may differ depending on the disk type (RAW, QCOW, etc.) [3,4]. Using the main memory which has disadvantages of high cost and resource constraints as a disk for virtual desktop, the memory capacity increases quickly and exceeds the total memory capacity as customer uses the virtual desktop and generates own data. This in-memory technique leads to an I/O bottleneck of the virtual desktop and the performance degradation.

In this paper, we propose a hybrid in-memory storage that integrates the main memory (e.g.

This work was supported by "Development of the Could-Based Data Sharing Environment" funded by Ministry of Land, Infrastructure and Transport in 2020. [No.19TLRP-B148994-02]

RAM) and disk storage (e.g. SSD, NVMe, etc.) in order to solve the shortcomings of the main memory capacity. By integrating main memory and disk storage into a single storage, we implement the hybrid in-memory storage, and provide a high-performance virtual machine for virtual desktop service and a large capacity of disk storage. Also, by applying hybrid in-memory storage to virtual desktop system, we intend to use it as a high-speed in-memory-based image storage disk.

Chapter 1 explains the overview of hybrid in-memory storage, Chapter 2 explains detailed structures and synchronous and asynchronous structures, Chapter 3 describes virtual desktop applications and experimental results, and Chapter 4 draws conclusions.

2. Overview of Hybrid in-memory storage

Hybrid in-memory storage is a virtual block storage device that uses a certain unit (gigabyte unit) of memory as a RAM disk and backs up data synchronously or asynchronously to a high-performance and non-volatile block device such as an SSD.

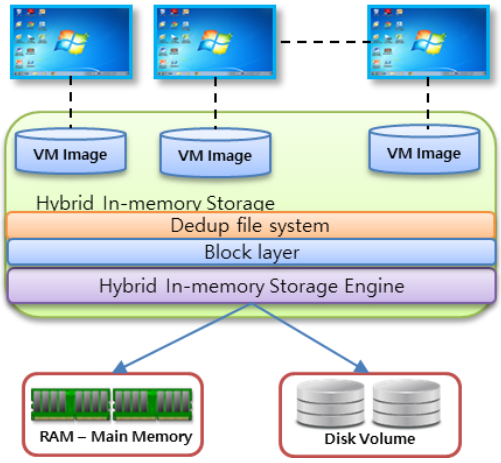


Fig. 1. Structure of hybrid in-memory storage.

Fig. 1 shows the structure of hybrid in-memory storage proposed in this paper. The entire system consists of a virtual machine, virtual image, hybrid in-memory storage, hybrid in-memory storage engine, main memory (RAM) and disk storage (SSD). Virtual machines create and operate virtual images (e.g. OS image, Disk image, etc.) on hybrid in-memory storage. The hybrid in-memory storage engine creates a single storage and combines main memory and disk storage to a single storage for hybrid in-memory storage. In addition, hybrid in-memory storage provides a standard block storage I/O interface, allowing existing virtual machines to be operated without modification.

Virtual disk access commands generated from the virtual machine are delivered to hybrid in-memory storage. Because the hybrid in-memory storage engine combines main memory and disk storage to create hybrid in-memory storage, hybrid in-memory storage engine handles disk access commands by selection of the integrated main memory and disk storage. In addition, hybrid in-memory storage can greatly reduce the actual virtual desktop image by using a file system through deduplication for limited memory operation.

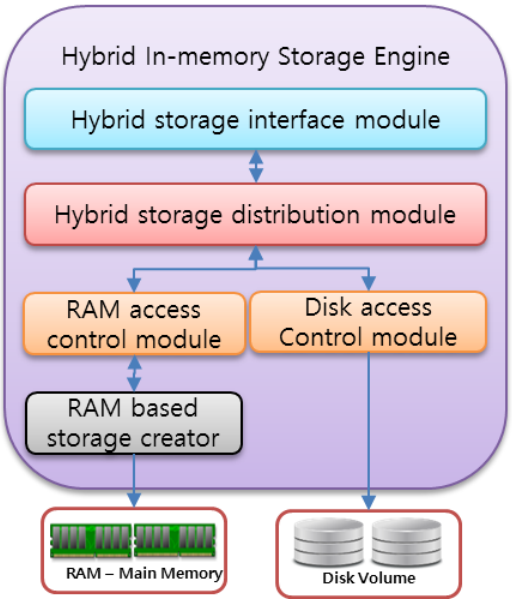


Fig. 2. Structure of hybrid in-memory storage engine.

Fig. 2 shows the structure of the hybrid in-memory storage engine in details. The hybrid in-memory storage engine consists of a hybrid storage interface module, a hybrid storage distribution module, RAM access (main memory) control module, RAM based storage creator, and disk access control module. The hybrid storage interface module provides an interface in a standard block storage format and receives virtual disk access commands generated

from virtual machines. The received command is passed to the hybrid storage distribution module. The hybrid storage distribution module determines whether to perform the command using main memory or disk storage according to the characteristics of the command, and transmits the command to RAM access control module or disk access control module.

The RAM access control module processes the disk access commands using the main memory as disk, which provides high speed access. The main memory storage generation module performs an actual read/write operation on the main memory that can be accessed in units of addresses with disk access commands transmitted in block units. Through this, the data of the virtual disk are stored in the main memory. The disk storage control module uses disk storage to process virtual disk access commands.

3. Operation of hybrid in-memory storage

Fig. 3 shows how hybrid in-memory storage initially operates. When the system starts, the data in the storage is restored to main memory (RAM disk) in the hybrid in-memory storage. Even during the restoration, the hybrid in-memory storage engine can continue the service without stopping. After restoration, the hybrid in-memory storage engine performs read/write(backup) operation. In synchronous mode, all data transferred to the hybrid in-memory storage engine is simultaneously stored in the RAM disk and SSD. When storing to both devices is completed, writing to the upper layer is completed. But, in the asynchronous mode, data transferred to the hybrid in-memory storage engine is only stored in the RAM disk. The kernel thread inside the hybrid in-memory storage engine stores data in RAM to the SSD in LRW (Least Recently Written) order.

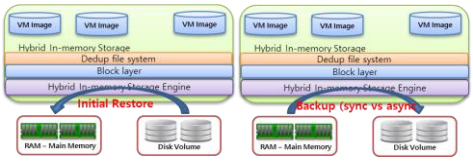


Fig. 3. The operation of hybrid in-memrory storage.

As shown in the **Fig. 4**, the capacity of hybrid in-memory storage can be set by the capacity of SSD. Then, the data within the RAM size uses the RAM disk, and the excess area is serviced directly from the SSD.

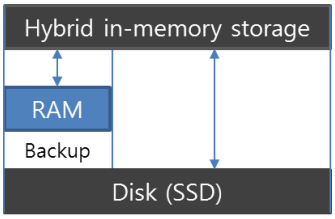


Fig. 4. The capacity of hybrid in-memrory storage.

Fig. 5 shows how to combine RAM disk and disk storage to create a hybrid in-memory storage. Hybrid in-memory storage provides a standard block storage format, mapping the RAM disk area at the front of the hybrid in-memory storage and disk (SSD) area at the rear.

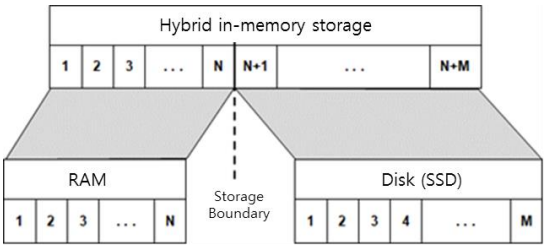


Fig. 5. The combining of hybrid in-memrory storage.

Block ID 1 to N of RAM are mapped to block ID 1 to N of hybrid in-memory storage. Block ID 1 to M of disk (SSD) are mapped to block ID N+1 to N+M of hybrid in-memory storage. And a storage boundary is set between the hybrid in-memory storage block ID N and N+1. When a disk access command is received, the block ID of the command is examined. If the block ID is less than or equal to the storage boundary, the command is transferred to the RAM access control module, and the command is processed using the high-speed main memory. If the block ID is larger than the storage boundary, the command is passed to the disk access control module so that it can be processed in a large amount of disk storage.

3.1 Synchronous Read/Write

Fig. 6 shows the synchronous read/write operation. Read operation is independent of the

SSD and are handled only in kernel memory in hybrid in-memory storage, internally. Intel's DMA Engine is used to transfer data from hybrid in-memory storage to user buffers. [5] In other words, the DMA Engine is executed without copying from CPU. In synchronous write (backup), writing is performed simultaneously to kernel memory and SSD. The requested write is completed only when the DMA copy of the kernel memory is completed and the writing of the SSD is also completed. All data movement is handled by DMA.

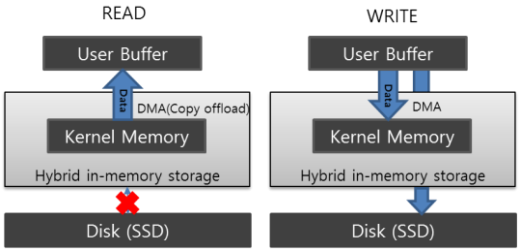


Fig. 6. The synchronous Read/Write operation.

3.2 Asynchronous Write

Fig. 7 shows the asynchronous write operation. In case of asynchronous write operation, all write operations are performed in kernel memory. In other words, when writing is completed only in kernel memory, the completion message is returned to the upper level. The written data in the memory is later written to the SSD. Even if the SSD's write speed is slow, the upper write request is not blocked, so the final write speed is similar to that of the RAM disk.

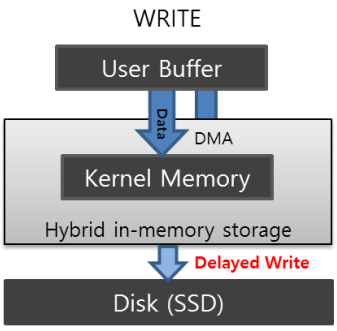


Fig. 7. The asynchronous Write operation

4. VDI Test & Results

The structure of a VDI system for applying hybrid in-memory storage is shown in figure 8. The VDI platform was configured using Intel(R) Xeon(R) CPU E5-2697 v3 @ 2.60GHz 64Core, CentOS 7.0, KVM-qemu hypervisor [6], and the virtual desktop operating system was configured with window10, and installed on hybrid in-memory storage.

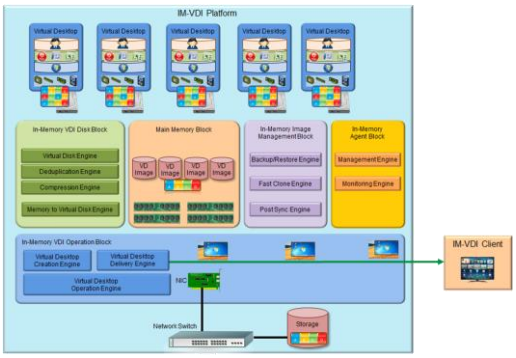


Fig. 8. VDI system for testing hybrud in-memory storage

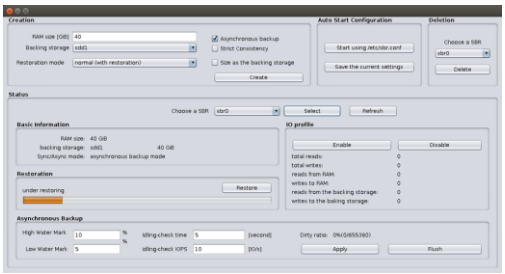


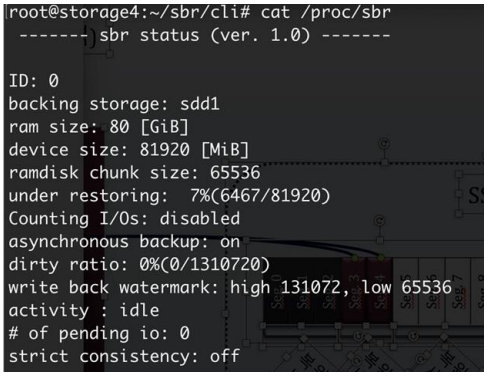
Fig. 9. VDI system for testing hybrud in-memory storage

Table 1. Testing enviorment for VDI system

Testing Environment	
CPU	Intel® Xeon® CPU E5-2697A 2.6GHz 64Core(16)
RAM	755 GB
Storage	2.6 TB (SSD)
OS	CentOS 7.2.1511
Kernel	3.10.0-327
Hypervisor	qemu-kvm
Network	1G Ethernet
No. of Node	2 (Management Node + Operating Node)

In order to configure hybrid in-memory storage in the VDI system, storage module was installed and a GUI tool was executed to set the storage as shown in **Fig. 9**. The developed

storage kernel module was installed in Linux kernel 3.10.0. Hybrid in-memory storage has 755G Memory, 2.6TB SSD and the size of OS image per virtual machine is 1600MB after deduplication. The summarized testing environment is shown in **Table 1**. The configured storage can be checked in proc file system of Linux as shown in **Fig. 10**.



```
root@storage4:~/sbr/cli# cat /proc/sbr
 ------- sbr status (ver. 1.0) -------

ID: 0
backing storage: sdd1
ram size: 80 [GiB]
device size: 81920 [MiB]
ramdisk chunk size: 65536
under restoring:  7%(6467/81920)
Counting I/Os: disabled
asynchronous backup: on
dirty ratio: 0%(0/1310720)
write back watermark: high 131072, low 65536
activity : idle
# of pending io: 0
strict consistency: off
```

Fig. 10. Proc file for hybrid in-memeory storage.

In the virtual desktop system, the most commonly used crystal disk mark for storage testing was installed and tested on the virtual desktop. [7] For Seq Q32T1, it represents the speed when reading and writing files in the queue in sequence. In other words, it refers to the speed when reading and writing 128KiB size with 32 instructions. In the case of the 4K Q32T1, since 4KB is used as a cluster in the NTFS system which is a common format method of the Windows operating system, the speed is measured while randomly processing the number of queues and threads, respectively. In the case of the crystal disk mark, Seq Q32T1 is regarded as the maximum value of read/write speed, and 4K Q32T1 is regarded the actual read/write speed. This experiment was tested using the benchmark tool (crystal disk mark), and is the result value measured through 10 periodic tests for each read/write operation.

Using the measured result values, the average value of the two benchmark results was obtained. The average value is shown in Table 1. As shown in **Table 2**, comparing with the average value, it can be nature that the two tested benchmark cases have significantly improved performance compared to general storage. But the more important point is the improvement of performance. The 4K Q32T1 showed better improvement of performance than Seq Q32T1, and the performance was improved by 42 times in reading operation and 6 times in writing operation. It was confirmed that the performance improvement was less in the case of writing operation than reading operation due to the characteristics of the Windows system and the asynchronous writing using DMA.

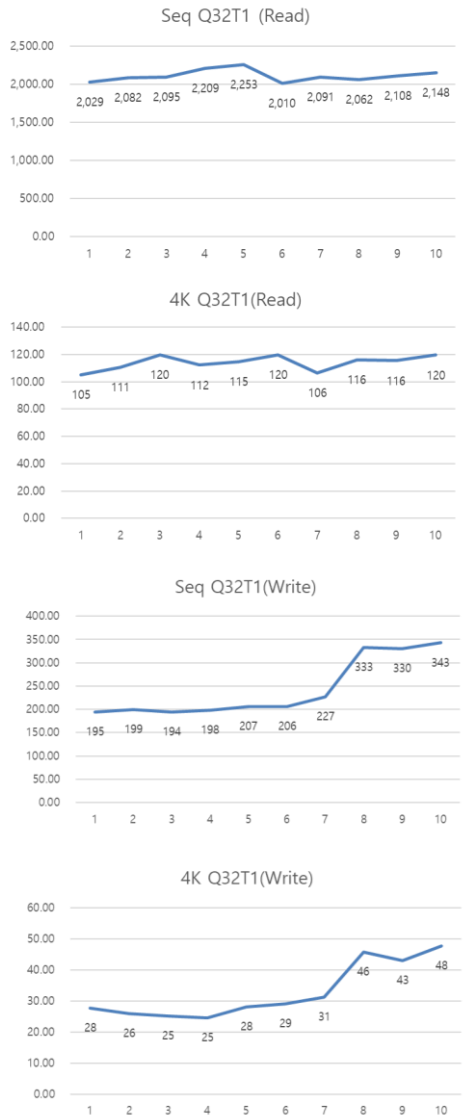


Fig. 11. The benchmark results. (Crystal disk Mark)

Table. 2 The average Speed (MB/S)of Read/write operation

	Disk based VDI		Ours	
	Read	**Write**	**Read**	**Write**
Seq Q32T1*	86.12	208.30	2,108.70	243.18
4K Q32T1**	2.70	4.98	114.02	32.82

* Seq Q32T1: Sequential (Block Size=128KiB) Read/Write [Queue:32/Thread:1]
** 4K Q32T1: Random 4KiB Read/Write [Queue:32/Thread:1]

4. Conclusions

In this white paper, we developed hybrid in-memory storage for the stability and high performance of virtual desktop infrastructure. By combining the characteristics of main memory and SSD, we overcome the shortcomings due to the volatility of the memory, and due to the high speed of the main memory (RAM), the performance of the virtual desktop using hybrid in-memory storage could be greatly satisfied as the test result.

For the presented virtual desktop service, the OS image that has the greatest impact on the speed of the virtual desktop is stored in main memory, allowing you to run high-speed virtual machines. When a virtual machine is deployed, users continue to create data while using their virtual desktops, and when this exceeds the main memory capacity, the data is stored on disk storage. Because user data is stored on disk storage, file access speed is slower than in-memory, but it is suitable for efficient operation and storage of large amounts of data. Therefore, OS images that are critical to the speed of the virtual machine are stored in main memory to support high performance, and large user data is stored in disk storage to overcome the capacity limitation of main memory. Also, due to the volatile nature of the main memory, it can reasonably cope with backup/recovery.

References

[1] Brian Madden, Gabe Knuth and Jack Madden, “The VDI Delusion” March. 20, 2012.

[2] Soo-Cheol Oh and Seong Woon Kim, “Design and implementation of high-performance virtual desktop system managing virtual desktop image in main memory,” *KIISE Transactions on Computing Practices*, Vol 22, No. 8, pp. 363-368, Aug, 2016.

[3] “QEMU Git tree - docs/interop/qcow2.txt,” *qemu.org*, May 29, 2018.

[4] "Understanding RAW Files Explained," *Luminous Landscape*. March, 2, 2011.

[5] Jonathan Corbet, Alessandro Rubini and Greg Kroah-Hartman, “Memory Mapping and DMA,” *Linux Device Drivers, 3rd Edition.*

[6] “QEMU System Emulation User’s Guide,” https://www.qemu.org/docs/master/system/index.

[7] https://crystalmark.info/en/software/crystaldiskmark/

An Efficient Test Case Selection and Change Detection Framework for Regression Testing

Israr Ghani, Wan M.N. Wan-Kadir, Adila Firdaus Binti Arbain and Omeed Kamal Khursheed
School of Computing, Faculty of Engineering, Universiti Technologi Malaysia (UTM)
Skudai, Johor Bahru, 81310, Malaysia
[israrghani@outlook.com, wnasir@utm.my, adilafirdaus@utm.my, omeed.kamal-1984@graduate.utm.my]
*Corresponding author: israrghani@outlook.com

Abstract

Regression testing is an essential but expensive activity in software development. Many RTS techniques, models, and frameworks have been presented in the past few decades, but these techniques, models, and frameworks have been implemented for small scale instead of medium (100-300 test cases) or large-scale (300+ test cases) software testing. This is because most existing RTS techniques, frameworks or models either require direct code dependency between tests and systems under test (SUT) or cannot be functional on medium or large scale systems with enough efficiency and time within existing resources. This study is an attempt to propose an improved regression testing framework to select better test suites and detect modified changes in modified test case; hence, reduce time and improve precision. Therefore, it can be said that the efficient test case selection and detection of the modified change in the test case for regression testing are still broadly open for further improvements.

Keywords: Regression testing, Software testing, Test case Selection, Test case selection framework, Quality assurance, Software testing approaches

1. Introduction

There are several vital reasons for projects to use regression testing. However, it may increase the time because of the evolution in iterative development and the efficiency of different software products on medium or large scale testing environment [1]. Several state-of-the-art TCS approaches are limited to evaluate time only where efficiency has been skipped to evaluate the TCS performance in regression testing [2]. Due to this reason, we need to identify the modified test case selection changes by using a TCS algorithm between the modified and actual test suite by adopting the RTS framework. Next, a dynamic algorithm requires to detect only the modified parts of the requested change to reduce time and improve efficiency on both medium test cases and large scale. Furthermore, test case selection and detection of newly added regression testing approach have several challenges such as (i) source code unavailability of applications, (ii) run-time application's testing activities alike binding, discovery along with product and services publishing into production, (iii) collaboration testing due to heterogeneous systems, (iv) [2, 3]. The efficient test case selection to reduce time and improve efficiency The number of test cases increases to deal with the modified changes, and limited testing time is also vital due to insufficient testing environments reported in several studies [4-6].

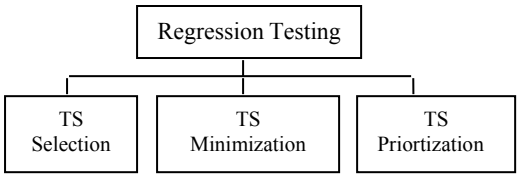


Fig. 1. Regression Testing

TC prioritization, minimization, and TC selection have been vital research with hundreds of papers investigating the different testing approaches in regression testing. Test case prioritization does not account for code modifications in SUT and also increases fault detection ability and ignoring the version of SUT. Therefore, our primary focus is on test case selection.

$$Time = (In\ Seconds)$$

$$Efficiency = \frac{TP\ (True\ Positive)}{TP\ (True\ Positive) + FN(False\ Nagetive)}$$

2. Regression Testing

Regression testing ensures that new faults will not be introduced into the modified code or requirement modifications in repetitive part of software testing. An effective test case selection frameworks approaches are generally used to improve the regression testing efficiently by selecting an efficient test suite without compromising the effectiveness [7, 8]. The outcome of effectiveness and the size of the test suite expressively increases by accepting all changes in testing, which may perhaps not be executed within allotted testing time. An efficient selection and detection of modified changes in test cases for regression testing are emerged to evaluate QoS parameters such as time and efficiency to make test suite according to the allocated time.

2.1. Test Case Selection Framework

The main purpose of this study is to propose an effective regression test case selection approach in terms of time performance and test suite precision to select an efficient and modified test suite. In otrder to address the above-mentioned problems, this study proposes test case selection and change detection frameworks to improve an efficient test case selection techniques, as well as to detect the newly added changes to reduce time and improve the precision of a regression testing. In **Fig. 2**, the selection process of the proposed framework where mapping rules such as (i) change history collection, (ii) calculate the number of modified changes, (iii) to select an efficient test suite from the number of total test cases modified in System Under Test (SUT) from test suite X to test suite Y.

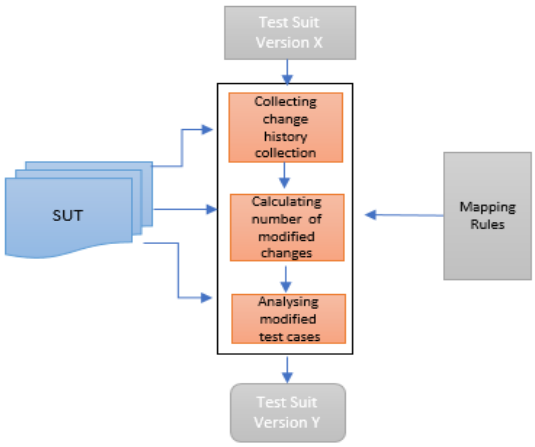


Fig. 2. Test Case Selection Framework

The framework should be well-structured in a planned manner to assist the researchers in accomplishing the research goals.

2.2. Change Detection Framework

The second part of the framework is showing the detection of the newly added change in the modified test case from test suite X where the change detection algorithm is applied for a small, medium, and large scale test suite, also filtered applied to exclude duplicate, fault detector, and already executed test cases according to the requirements to get the updated, and an efficient version of modified test suite Z. The standard interface, database, and API related format based formats and WSDL files will be used for this propose. It also derives the relationships among the different components. The proposed frameworks for test case selection and change detector framework can reduce time and improve precision by applying the mentioned rules. To

avoid repetition, for expected results in test case selection for small, medium, and large scale test suits, proposed frameworks are more efficient in terms of time and precision by dividing into test case selection, and change detection approaches.

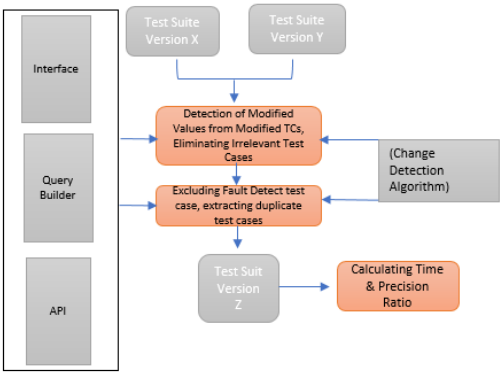


Fig. 3. Change Detection Framework

The main purpose of this change detection framework and algorithm is to achieve the performance of efficiency of the proposed approach in terms of time in improving the regression testing of software.

```
1: Test case id: Tc1
2Inputs
3: Aselectioset[]: Total test cases in suit1 query
4: A1[]: Modified Test Cases in suit2 query //Addition in the this algorithm
5: A2[]:
6:output
7: {SUT: publictestMaximumValue(int max)
8: Max= total number of test cases
9: if(Max==A)
10: Datetime dt = new DateTime(hour,mint,sec) //New dateTime = Current DateTime
11: A2= A-A1 // Addition of this calculation
//12: while(dt.getdatetime() < new dateTime) (Maral Azizi ,2019) //calculation of overall testing time
12: while(dt.getdatetime() < new dateTime) // Calculation of detected modified test case time only
13: SUT.valueprinttime(hour,mint,sec)}
14: return SelectionSet[];
15: return A1[]
16: return SUT[]
17: retun T[]
18: end procedure
```

Fig. 4. Pseudo code of Change Detector Algorithm

3. Conclusions

The aim of enhanced regression test case selection and detection of the modified change approach is to select a test suit that determines quality parameters such as time and efficiency are important to improve regression testing techniques. For future research, an enhanced regression testing approach needs more concentration for performance, integrity, security, and inconsistencies of regression testing.

References

[1] Yamuç, A., et al. Solving test suite reduction problem using greedy and genetic algorithms. in 2017 9th International Conference on Electronics, Computers and Artificial Intelligence (ECAI). 2017. IEEE.

[2] Agrawal, A.P. and A. Kaur, A comprehensive comparison of ant colony and hybrid particle swarm optimization algorithms through test case selection, in Data engineering and intelligent computing. 2018, Springer. p. 397-405.

[3] Agrawal, A.P., A. Choudhary, and A. Kaur, An Effective Regression Test Case Selection Using Hybrid Whale Optimization Algorithm. International Journal of Distributed Systems and Technologies (IJDST), 2020. 11(1): p. 53-67.

[4] Arcuri, A. RESTful API automated test case generation. in 2017 IEEE International Conference on Software Quality, Reliability and Security (QRS). 2017. IEEE.

[5] Alagöz, I., T. Herpel, and R. German. A selection method for black box regression testing with a statistically defined quality level. in 2017 IEEE International Conference on Software Testing, Verification and Validation (ICST). 2017. IEEE.

[6] Anabalon, D., et al. Test reduction for easing web service integration. in Simposio Argentino de Ingeniería de Software (ASSE 2015)-JAIIO 44 (Rosario, 2015). 2015.

[7] Silhavy, R., P. Silhavy, and Z. Prokopova, Analysis and selection of a regression model for the Use Case Points method using a stepwise approach. Journal of Systems and Software, 2017. 125: p. 1-14.

[8] Choi, W., et al. DetReduce: minimizing Android GUI test suites for regression testing. in 2018 IEEE/ACM 40th International Conference on Software Engineering (ICSE). 2018. IEEE.

Product Identification Technique Using Deep Learning in Smart Factory

Yewon Choi* and Jongeon Byeon and Nahun Lee** and Sanghwa Lee** and Hayong Jeon*****
*Management Information System Department, Kyungnam University
**Computer Science Department, Kyungnam University
***Division of Mathematics, Science, and Computer(MSC), Kyung-Nam University
Changwon, South Korea
[e-mail: wony9802@naver.com, plu125@naver.com, elen02@naver.com]
*Corresponding author: Nahun Lee

Abstract

This paper proposes a technique that can discriminate both genuine and defective products of products produced in smart factories using a convolutional neural network (CNN) layer. One of the deep learning techniques. A specific area of a product is detected and extracted by a camera using a learning model. And a CNN model is applied to train it to determine the product taken by the camera. By applying the technique implemented in this paper to a smart factory and discriminating the product in real time. It is possible to improve the production environment and increase the efficiency of production by reducing the discrimination time.

***Keywords*:** Deep Learning, Convolutional Neural Network(CNN), Smart Factory

1. Introduction

With the recent advent of the 4th Industrial Revolution. The governments of major countries have made many changes in society as a result of various efforts to provide differentiated policy support. And smart factories can be seen as the core technology of the manufacturing industry as the core technology of this change. A smart factory is a factory that has built a new system. That system is automatically producing through simulation of the machine itself through the fusion of ICT (Information Communication Technology) and the manufacturing industry [1,2].

According to Markets & Markets, the global smart factory (manufacturing) market size is expected to grow by 9.3% annually until 2022, forming a market size of 20.054.2 billion dollars, and the Korean market size is 783 billion dollars in 2020 and 2022. It is expected to reach $12.76 billion until 2010, with an annual growth rate of 12.2%, the second fastest growth rate in Asia after China [3].

Before the development of the smart factory, it was not possible to know whether the corresponding data was used for prediction. Also analysis in the process of collecting necessary data, and organic management was difficult. When these problems occurred, they were resolved through the analysis of skilled technicians. But they got some mistakes that were frequently made. Like time and cost to respond were incurred. That cause the decrease in the number of skilled workers to the aging of the current technicians and wage problems. Accordingly, this paper proposes a method of changing the existing management system method to build a smart factory environment. And applying and utilizing the CNN technique as the center of necessary data.

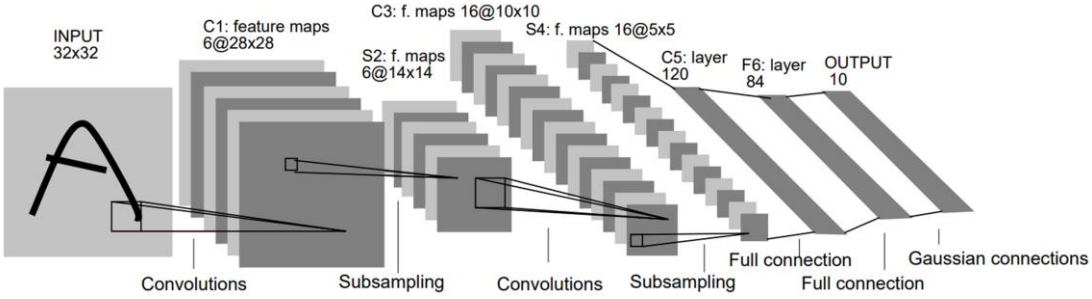


Fig. 1. The basic structure of CNN [5]

2. Related Research

Deep learning is an advanced version of an artificial neural network during machine learning, which is a method of expressing and learning gradual meaning in successive layers. In deep learning, the basic layered model is called a neural network model, and the word neural network is a name that was inspired by understanding human brain structure. In deep learning, the details of processing input data in a layer are stored in the weight of the layer consisting of numbers. Technically speaking, the transformation that occurs in a specific layer is expressed as a function that takes the weight of that layer as a parameter. In this form, learning means finding weight values in all layers of the neural network to map a given input to the correct target. However, some deep neural networks have tens of millions of parameters. Neural networks take weights as a parameter, and to control the output. We need to measure how far the output deviates from what we expect. This is the loss function or objective function. To measure how well the neural network predicted. For example, the difference between the predicted value of the neural network and the value expected from the output of the neural network is calculated as a score.

Convolutional Neural Network (CNN) is a method that is used to extract image features and process texture information. It has already overwhelmed the existing algorithms in 2012 ILSVRC (Imagenet Large Scale Visual Recognition Challenge) Performance was verified [4]. The basic structure of CNN is consist of feature extractor and classifier part, as shown in **Fig. 1**, extractor contain convolution layer and pooling layer. Classifier include dense layer and drop layer.

Convolution is a value obtained by multiplying adjacent pixels of output data by a convolution filter, and the influence of adjacent pixels is the output data.

3. Implementation and Results

3.1 Schematic of The Overall System Structure

Using the method proposed in this paper, we implement it based on **Fig. 2**, which is a schematic diagram of the overall system structure, It uses the most basic structure among CNN techniques.

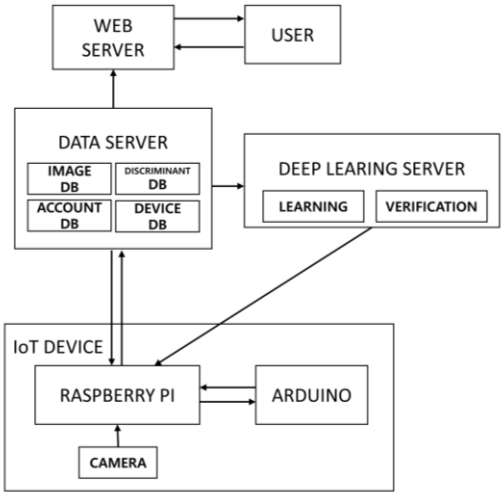


Fig. 2. Schematic of the overall system structure

3.2 Analysis Process

In image processing, in order to derive an optimized discrimination result for a moving product, light from outside is blocked and a

constant internal lighting brightness is set as a reference. After binarizing the image captured in a specific area, the contour is detected to determine the exact location information of the product. When reaching the designated location of the input product, the product and a part of the surrounding area are photographed to extract optimized data and obtain a uniform image. As for the input data, as shown in **Fig. 3**, existing images and captured images are used, and the image used for learning is changed to a size of 120*120 and preprocessed.

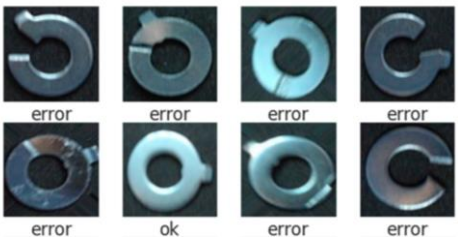


Fig. 3. Part of the image used for training

According to the model created by the initial training, the corresponding image outputs the result of discrimination and similarity. During the execution process, the sequence of images used for training is randomly used. So a part of the image is extracted and displayed on the screen as shown in **Fig. 4**. And the layer type and output form of the deep learning model to be used for training are displayed.

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 60, 60, 16)	448
conv2d_1 (Conv2D)	(None, 30, 30, 32)	4640
max_pooling2d (MaxPooling2D)	(None, 15, 15, 32)	0
conv2d_2 (Conv2D)	(None, 8, 8, 64)	18496
max_pooling2d_1 (MaxPooling2	(None, 4, 4, 64)	0
conv2d_3 (Conv2D)	(None, 2, 2, 128)	73856
max_pooling2d_2 (MaxPooling2	(None, 1, 1, 128)	0
flatten (Flatten)	(None, 128)	0
dropout (Dropout)	(None, 128)	0
dense (Dense)	(None, 256)	33024
dropout_1 (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 210)	53970

Total params: 184,434
Trainable params: 184,434
Non-trainable params: 0

Fig. 4. Model structure

3.3 Analysis Result

The trained model created as a result of the analysis has an accuracy of about 90%, trains 100 times. And consumes an average of 3 to 5 seconds per epoch, which consumes a total of 5 to 8 minutes. In addition, the verification accuracy and loss rate are displayed each time the corresponding round is over during learning. When learning is completed, the accuracy change and loss rate change for each round can be checked with a graph as shown in **Fig. 5** and **Fig. 6**.

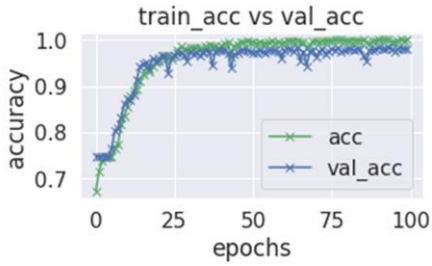


Fig. 5. Accuracy rate graph between training

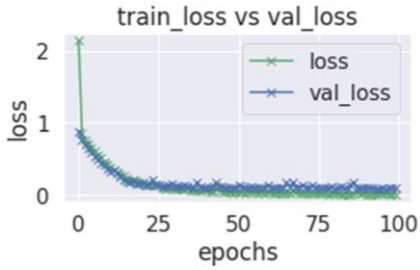


Fig. 6. Loss rate graph between training

4. Conclusions and Expected Effect

Using the modified technique proposed in this paper, The product is automatically identified by the machine for defective products. And the results are connected by wireless communication, allowing free linkage between the preceding and following processes. In addition, by using the learned data and stored images, the accuracy of the model is increased and the learning time is shortened. This quickly finds defects. Since the technique proposed in this study can be applied to various products. It can be used in various fields.

References

[1] Krizhevsky Alex, Ilya Sutskever, and Geoffrey E.Hinton, "Imagenet classification with deep convolutional neural network", *Advances in neural information processing systems*, 2012.

[2] Cha Seok Geun, "The trend and system structure of standardizing smart factory using IoT/M2M", *The Journal of The Korean Institute of Communication Sciences 32.5 (2015)*: 36-41.

[3] Automotive Cyber Security Market Research Report - Global Forecast to 2025 - Cumulative Impact of COVID-19, *MarketsandMarkets*, 2020.

[4] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, et al, "Imagenet large scale visual recognitionchallenge," *International Journal of Computer Vision*, Vol 15, No 3, pp. 211-252, 2015.

[5] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-based learning applied to document recognition", *Proceedings of the IEEE*, vol. 86, issue 11, pp. 2278-2324, November, 1998.

Development of an Air Quality Monitoring Device for Accommodation Facilities based on Service Design Methods

Kwanpyo Choi*, Taekheon Ahn*, Semi Park*, Mihee Han, Ilsoon Bang*****
* Grib Inc. (kpchoi@grib.co.kr, tekern@grib.co.kr, semi0095@grib.co.kr)
** P&D Design (pnddesign@hanmail.net)
*** ELT Sensor (isbang@eltsensor.co.kr)

Abstract

In this paper, we present a harmful gas detection solution developed to provide smart accommodation services. The smart accommodation services aim to make customers more comfortable in lodging facilities by automatically and intelligently control smart equipment. We used the service design techniques to drive smart accommodation services that customers want unconsciously and developed an air quality monitoring device to implement them. In the device, we used diverse environmental sensors such as CO2, CO, VOC, particulate matter as well as the temperature and humidity sensor. To develop more sophisticated smart lodging services, we conducted various evaluation tests to understand the operation characteristics of individual sensors.

***Keywords*:** Air Quality, Monitoring, IoT, Accommodation Facilities, Service Design, Fire

1. Introduction

With the spread of COVID-19, a paradigm shift is appearing in the tourism industry. People are flocking to domestic trips instead of overseas trips, and trips that enjoy leisure is taking its place rather than trips full of various schedules. In fact, wellness tourism is growing 50% faster than the overall tourism industry [1]. In addition, interest in small-scale facilities such as pensions that can minimize contact with people rather than large accommodation facilities such as hotels and resorts is increasing significantly.

However, people think that small & medium-sized accommodation facilities are somewhat inconvenient and unsafe for several reasons, even though the room charge is cheap. In particular, the accident due to toxic gas from boiler at a pension in Gangwon-do in December 2018 has raised concerns of many. After that, accordingly, operators of small & medium-sized accommodation facilities are striving to provide more comfort-able and safe travel experiences based on digital technol-ogies including IoT.

In fact, some US companies such as Pure and Delos deliver services to provide a pleasant experience in accommodation[2]. However, due to its high price, only large hotel chains such as Hilton, Hyatt, and Marriott are intro-ducing it to a limited extent. Therefore, we decided to develop an in-room air quality management service that can be used casually in small & medium-sized accommodation facilities.

First, we decided to use service design techniques to understand the exact needs of the

This research was supported by the MOTIE(Ministry of Trade, Industry and Energy), Korea, under the Industrial Technology Innovation Business support program (No.20004393, Development of platform design technology for fire and air quality detection service for accommodation and joint residential facilities) supervised by the KEIT(Korea Evaluation Institute of Industrial Technology). We'd like to special thank Prof. Hak-Yong Kim for his advice and guidance.

market and then to develop such a device that detects various types of harmful gases to meet the needs. The device can measure CO2 and CO, VOCs and particulate matter as well as temperature, humidity, and smoke.

We have already developed and reviewed several service scenarios especially in terms of in-room air quality management services. However, we will not mention them in this paper any more since the services must be based on accurate and stable sensor data. We rather focus on the performance test results of some sensors.

This paper is organized as follows: First, we introduce the background of developing a harmful gas detection device in Chapter 2. We explain the structure and operation method of the harmful gas detection device and present the operating characteristics of sensors used in harmful gas detection devices in Chapter 3 and 4, respectively. In Chapter 5, finally, we conclude this paper with an introduction to the future works.

2. Device Development based on Service Design Methods

Various smart devices for small & medium-sized accommodations are already on the market. For example, a smart fire detector is typical which informs the manager or fire department of fire outbreak. In addition to this, many smart switches and plugs are also used to automatically turn on and off garden lights or turn on and off the power of insect repellent when it gets dark. Smart gas valves, smart door locks, and intelligent CCTV are also commonly used devices to improve the safety and convenience of accommodation services.

However, most of these devices are intended to improve the convenience of an operator or manager of a lodging facility rather than a safe and comfortable stay of the customer. Of course, it is important to promote operational efficiency, but it is more important to think about and provide services from the perspective of customers in the future. Therefore, we used service design techniques to understand the hidden needs for the services customers want for accommodation.

Fig. 1. Service Design Workshop for Customer-Oriented Lodging Service

We held a service design workshop to derive accommodation services from the perspective of customers, and a total of 16 people partici-pated. In the beginning, we tried to derive several ideas and important keywords for lodging services through brainstorming. Based on this keyword, over 200 ideas were categor-ized, and among them, services that were considered important from a customer's point of view were selected. For example, windows are automatically opened or closed according to the air condition in the room or announce-ments are made in the guest's native language in case of fire.

For some of the selected services, we made an MVP(minimum viable product) using ESP32 and reviewed the possibility of realization and commercialization of the service [3]. In this process, we decided that it was important to accurately measure the air condition of guest rooms and further predict changes in air condition. As a result, we decided to develop a device that measures the temperature and humidity in guest rooms, as well as CO, CO2, VOCs, smoke, and fine dust, and to under-stand their exact operating characteristics. We also decided to develop a technology that predicts changes in environmental indicators by analyzing datum obtained from the device.

3. Architecture of the Air Quality Monitoring Device and Its Operation

Fig. 2 shows the architecture of the air quality monitoring device. As shown in the Fig., the device contains five sensors, including a temperature and humidity sensor, CO, CO2, TVOCs, and PM 2.5. The temperature and humidity data are used to maintain the proper

temperature and humidity in the room, and CO, CO2, TVOCs, and PM 2.5 are used to maintain a safe and comfortable air condition.

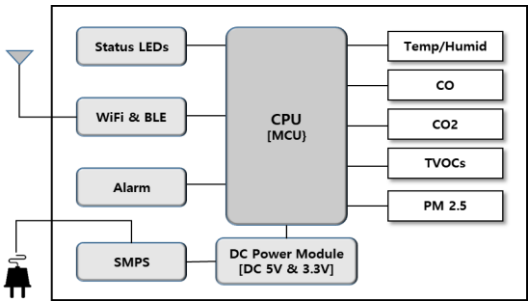


Fig. 2. Architecture of the Air Quality Monitoring Device

We use MXH2101 as the temperature and humidity sensor which can measure tempera-ture in the range of -40°C ~125°C and humidity between 20~80% [4]. As the CO sensor, we use CO-S20-3V, an electrochemical type sensor, which measures CO in the range of 0~1000ppm [5]. The CO2 sensor is also an electrochemical type and provides a measure-ment range of 400~5,000ppm. The TVOC sensor is MS-VOC01-5V and the fine dust sensor, SM-LS-01B, can measure up to 1000μg /m3 [6].

In addition to five sensors, this device includes a display unit and an alarm unit to display the air condition, a communication module that transmits processed sensor data to an external service system, and a CPU which stores and processes the data generated by the sensor. ESP32 is used as a communication module since it includes both Wi-Fi and low-power Bluetooth to connect to a smartphone as well as a Wi-Fi router. In addition, displays, alarms, and power supplies use conventional Mono LCD, Piezo Buzzer, and SMPS-based power supplies.

Fig. 3 shows the operating sequence of the air quality measuring device. This device operates in Station Mode and can send and receive data to and from local devices or external networks through a Wi-Fi router. When a client is connected, the sensor data is transmitted every 1 second in the UDP format, and the sensor data is transmitted to the DB server every 30 seconds. This device can also be connected to a user's smartphone using Bluetooth, but a description of this part will be left out in this paper.

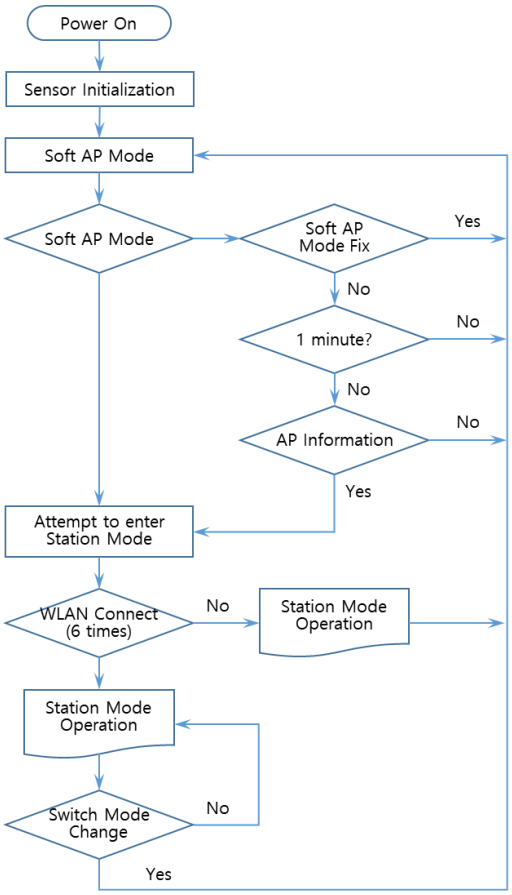


Fig. 3. Operation Flow chart of the Air Quality Monitoring Device

4. Analysis of operational characteristics of sensors

When the test device (DUT) is connected to the Wi-Fi router, it transmits sensor data to the DB server at regular intervals. Sensor data can be usefully used in services only when the response rate to environmental changes or the duration of state change is accurately known. Accordingly, we conducted various operation tests on the sensors. In this paper, we would like to introduce four typical features.

First, the reaction time was tested for the CO sensor. According to the test result, it took 13 seconds to reach T1/e and 27 seconds to reach T90. It responds very quickly to changes in the CO value, so it is expected to be able to detect a fire quickly.

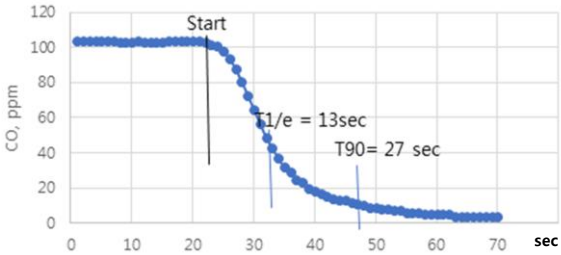


Fig. 4. Reaction time test of CO Sensor

The CO2 sensor was also somewhat slower than the CO sensor, but showed relatively fast response time characteristics. T1/e was 70 seconds and T90 was 100 seconds. In addition, the measurement accuracy was also evaluated very stably, and the measurement error rate in various situations was less than 4%.

Input (ppm)	Output (ppm)	Error Rate
500	504	0.86%
1,000	1,034	3.36%
2,000	2,058	2.89%
4,000	4,137	3.43%

Table. 1. Accuracy test result of the CO2 sensor

The **Fig. 5** below shows the response test result of the VOC sensor. After applying 19.4ppm C_4H_8 gas at 25°C, the reactivity of the sensor output was tested. It was confirmed that the saturation state was reached 120 seconds after the gas was injected, and it was confirmed that the state corresponding to 90% of the maximum value continued for 180 seconds or more.

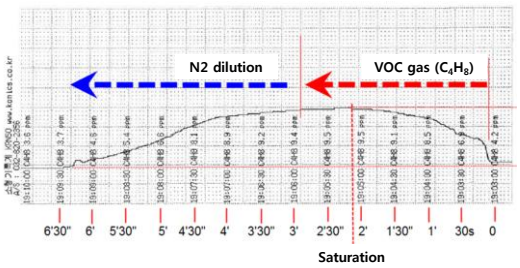


Fig. 5. Reaction Time Test of VOC sensors

Fig. 6 shows the results of the reaction test of the smoke sensor. After applying cigarette smoke to a smoke sensor installed in a small chamber at 25°C, the sensor output is saturated after about 50 seconds. After that, it shows that this state lasts for about 100 seconds or more. The smoke sensor output value in this state was 9.8ppm, which can be used as an indicator for sounding a fire alarm or operating a fan and air purifier.

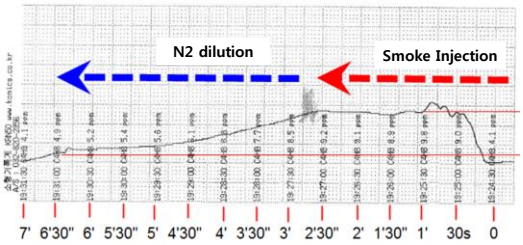


Fig. 6. Reaction Time Test of Smoke sensors

5. Conclusions and Future Work

In this paper, we introduced the process of deriving various services for small & medium-sized accommodation facilities using service design techniques and the air quality measurement device required for these services. In addition, the structure and operation method of the air quality measuring device were explained, and the operating characteristics of the five sensors used in the device were introduced. The reaction time of the sensor to environmental changes or the time it takes to reach saturation can be useful in triggering or deactivating each service. Also, it will be useful when handling exceptions for very short fluctuations in sensor values. In the future, we will develop a method how to start the service more accurately and quickly through big data analysis on the change pattern of sensor data.

References

[1] Betsy Goldbergj, "7 Big Trends in Wellness Tourism to Experience Now," Virtuoso, Aug.22, 2016.

[2] Global Wellness Institute, Global Wellness Tourism Economy, Nov. 2018.

[3] ESP32 Series datasheet, https://www.espressif.com/sites/default/files/documentation/esp32_datasheet_en.pdf

[4] MXH2101 Datasheet, https://www.alldatasheet.com/view.jsp?Searchword=MXH2101

[5] CO-S20-3V datasheet (in Korean), http://product.ic114.com/PDF/C/DS_CO-S20-3V.pdf

[6] SM-LS-01A, http://www.nids.co.kr/product2.php

Fingerprint-based Beam Selection and Cooperation for Millimeter-wave UAV Communication

Huaping Liu[1], Hyeonsung Kim[2], Intae Hwang[2], and Sangmi Moon[2]
[1]School of Electrical Engineering and Computer Science, Oregon State University
Corvallis, USA
[e-mail: huaping.liu@oregonstate.edu]
[2]Department of Electronic Engineering, Chonnam National University
Gwangju, Republic of Korea
[e-mail: hsk940127@naver.com, and hit@jnu.ac.kr, and msm0804@jnu.ac.kr]
*Corresponding author: Sangmi Moon

Abstract

In this paper, we propose a fingerprint-based beam selection and cooperation scheme for millimeter-wave (mmWave) unmanned aerial vehicle (UAV) communication. The proposed scheme constructs an offline fingerprint database for beam selection and performs online beam cooperation. In the offline phase, the best beam index from serving cells and the interference beam indexes from neighboring cells are stored. In the online phase, the best beams and interference beams are determined using the information from the fingerprint database instead of an exhaustive search, and the beam cooperation is performed to improve the signal-to-interference-plus-noise ratio for aerial user equipments. System-level simulations are performed to assess the UAV effect based on the 3GPP new radio mmWave and UAV channel models. Simulation results show that the proposed beam selection scheme can reduce the beam sweeping overhead and inter-cell interference.

***Keywords*:** beam cooperation, beam selection, fingerprint, mmWave, UAV

1. Introduction

In the past few years, unmanned aerial vehicles (UAVs) have been increasingly employed in civilian applications, such as for aerial surveillance, traffic control, photography, package delivery, and communication relaying. According to a report from the Federal Aviation Administration (FAA), the number of drones will more than double to over 2.4 million by 2022 [1].

Millimeter-wave (mmWave) communication, which utilizes the wide available bandwidth above 28 GHz, is a promising technology to achieve high-rate UAV communication [2]. While mmWave communication has been extensively investigated for 5G-and-beyond cellular systems, its application to cellular-connected UAV systems faces both new opportunities and challenges. On the one hand, as mmWave signals are vulnerable to blockage, the line-of sight (LoS)-dominating UAV– base station (BS) channels offer the most favorable

This research was supported by the Ministry of Science and ICT (MSIT), Korea, under the Information Technology Research Center (ITRC) support program (IITP-2020-2016-0-00314) supervised by the Institute for Information & communications Technology Planning & Evaluation (IITP). It was also supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT: Ministry of Science and ICT) (2018R1A2B6002255 and 2020R1I1A1A01073948).

channel conditions for the practical application of mmWave communication. On the other hand, the high UAV altitude and mobility require the development of efficient mmWave beamforming for 3D mmWave UAV–BS channels. The beamforming technology can determine the best beam direction formed by multiple antenna elements for user equipment (UE) to maximize the transmission rate and improve the energy efficiency.

In this paper, we propose a fingerprint-based beam selection and cooperation scheme for mmWave UAV communication. The proposed scheme constructs an offline fingerprint database for beam selection and performs online beam cooperation. In the offline phase, the best beam index from serving cells and the interference beam indexes from neighboring cells are stored. In the online phase, the best beams and interference beams are determined using the information from the fingerprint database instead of an exhaustive search, and beam cooperation is performed to improve the signal-to-interference -plus-noise ratio (SINR) for aerial UEs.

2. Fingerprint-based Beam Selection and Cooperation

2.1 Fingerprint Database Construction

The measurement is performed in each cell by assigning the terrestrial and aerial UEs to each fingerprint spot, and the measurement content includes the fingerprint position, service cell ID and corresponding optimal beam ID. If the UE is an aerial UE, the interference cell IDs and the corresponding strongest interference beam IDs also need to be measured. Then, the measurement result is saved in the corresponding fingerprint database, i.e., a set of all the fingerprint information within a cell coverage area. Each fingerprint database is divided into two parts, namely the terrestrial and aerial fingerprint datasets. **Table 1** shows the fingerprint database of the cell b, in which $P_a(a=1,2,...,A)$ and $P_t(t=A+1,A+2,...,A+T)$ are the aerial and terrestrial fingerprint position, respectively. $I\_ID_i(i=1,2)$ is an interference cell ID for a fingerprint position, b_t^* is the optimal beam ID for P_t, and b_a^* and $\tilde{b}_a^i$ are the

Table 1. Example of fingerprint for the cell b

Aerial UE						
P_1		P_2		…	P_A	
b_1^*		b_2^*		…	b_A^*	
I_ID_1	$\tilde{b}_1^1$	I_ID_1	$\tilde{b}_2^1$		I_ID_1	$\tilde{b}_A^1$
I_ID_2	$\tilde{b}_1^2$	I_ID_2	$\tilde{b}_2^2$		I_ID_2	$\tilde{b}_A^2$
Terrestrial UE						
P_{A+1}		P_{A+2}		…	P_{A+T}	
b_{A+1}^*		b_{A+2}^*		…	b_{A+T}^*	

optimal beam ID and the strongest interference beam ID, respectively, corresponding to the interference cell i for P_a. Note that the number of interference cells is two ($i=1,2$) because we consider the intra-site joint transmission (JT) for aerial UEs, as will be discussed in the following subsections.

During the construction phase of the offline fingerprint database, the UE of each fingerprint position performs the beam selection. In the 5G standard, the main procedure for beam selection is beam sweeping, which transmits a set of predefined beams to cover the spatial area. Among the predefined spatial directions, beam sweeping exhaustively searches for the optimal beam transmitted by the service cell and the strongest interference beams caused by the adjacent cells. Here, we assume that all the beams generated by the same cell are orthogonal. Furthermore, the same beam serves multiple UEs through a time-division multiplexing mechanism to avoid intra-cell interference.

2.2 Beam Cooperation

In the online phase, the BS can further determine to which fingerprint dataset of the service cell of the UE the matched fingerprint position belongs. If the matched fingerprint position is in the aerial fingerprint position, the UE can be identified as an aerial UE; otherwise, it is a terrestrial UE. When a UE is a terrestrial UE, the serving cell only needs to transmit the best beam to the UE according to the beam ID corresponding the matched fingerprint position. As the aerial UEs experience LoS propagation conditions for more cells with higher probability, in comparison with terrestrial UEs, the aerial UEs receive interference from more cells in the downlink. Therefore, JT is applied to improve the

performance of an aerial UE. In this study, multiple cells belonging to the same site are coordinated, and data are jointly transmitted to the UEs. As intra-site JT-coordinated multi-point is already supported in 3GPP standardization, enhancements are not required. With beam cooperation for aerial UE, the interference from the neighboring cells is reduced because a neighboring cell is converted from the interference signal to the desired signal.

3. Simulation Results

System-level simulations are performed to assess the UAV effect based on the 3GPP new radio mmWave and UAV channel models [3,4]. The following schemes are simulated for comparison:
(1) Random beam selection (R-BS): each UE is served with only its serving cell, and the beams are selected randomly from among all the beams.
(2) Exhaustive search beam selection (ES-BS): each UE is served with only its serving cell, and the beams are selected via an exhaustive search.
(3) Beam selection with single-cell transmission (SC-BS): each UE is served with only its serving cell, and fingerprint-based beam selection is used.
(4) Beam selection with JT (JT-BS): the same as scheme 3, but intra-site JT is used for aerial UEs.

Fig. 1 plots the CDF of the SINR for schemes 1 to 4, where the aerial UE ratio is 50% (5 aerial UEs per cell). The proposed beam selection (JT-BS) can achieve a higher SINR than the other schemes.

Fig. 2 plots the CDF of the SINR for schemes 3 and 4 in cellular-connected UAV communication when the aerial UE ratios are 7.1% (case 3) and 50% (case 5). With the increase in the number of aerial UEs, the inter-cell beam interference increases, and the value of SINR in scheme 4 is significantly improved compared with that in scheme 3.

4. Conclusion

In this paper, we proposed a fingerprint-based beam selection and cooperation scheme for mmWave UAV communication. The proposed scheme constructs an offline fingerprint database for beam selection and performs online beam cooperation. The simulation results showed that the proposed beam selection scheme could reduce the beam sweeping overhead and inter-cell interference.

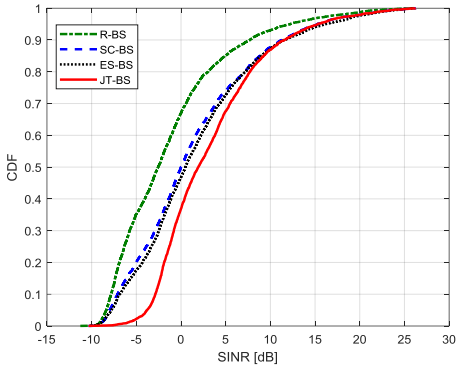


Fig. 1. CDF of the SINR for scheme (1)-(4)

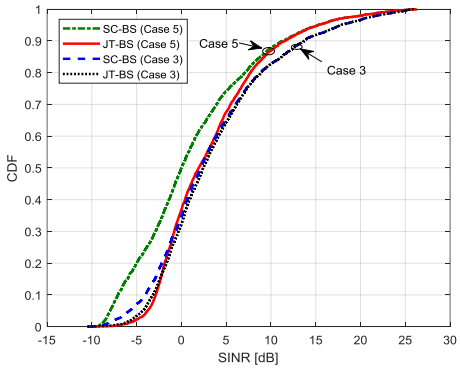


Fig. 2. CDF of the SINR for scheme 3 and 4

References

[1] H. Price, "Federal Aviation Administration (FAA) Forecast Fiscal Years 2017–2038." Feb. 2019.
[2] Z. Xiao, P. Xia, and X.-G. Xia, "Enabling UAV cellular with millimeter-wave communication: potentials and approaches," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 66–73, May 2016.
[3] 3GPP Technical Report 36.777, "Technical specification group radio access network; Study on enhanced LTE support for aerial vehicles (Release 15)," V15.0.0, Dec. 2017.
[4] 3GPP Technical Report 38.901, "Study on channel model for frequencies from 0.5 to 100 GHz (Release 16)," V16.1.0, Dec. 2019.

Bit Swarm Enhancements based Wireless Sensor Networks

Hawezi Omeed kamal khorsheed[1*], MD Asri Bin Ngadi, Abdul Hanan Abdullah[1], Israr Ghani
[1]Faculty of Computer Science, University Technology Malaysia (UTM), Malaysia
Skudai, Johor Bahru, 81310, Malaysia
Omeed.kamal-1984@graduate.utm.my, hanan@utm.my, dr.asri@utm.my, israrghani@outlook.com
Corresponding author :omeed kamal

Abstract

Wireless Sensor Networks (WSNs) are self-organizing, self-esteem and integrated sensors networks. These networks contain tinny sensor nodes with low power, less computational and storage capabilities. The sensor nodes are sensing, monitoring and communicating the data and then forward the data for further processing. The scope and structure of these networks have enormous impacts on system lifetime. In this paper, atom swarm algorithm is used as a procedure to improve the coverage and location of sensor nodes in the network. A framework proposes based on particle swarm. The proposed model is used as a piece for covering the sensor nodes transmission with respect to the scope of the given field. The centers are tolerating and being able to change their position.

***Keywords*:** WSN, practical swarm optimization, MEMS micro- electro- mechanism, maximum-support- path, maximum- breach- path.

1. Introduction

With advancements of new wireless technologies and noticeable improvements of Micro Electro Mechanism (MEMS), offer the fabrication, minimization based on low power solutions. These integrated systems are used for estimation, revelation, and data correspondence in encompassing condition. Wireless Sensor Networks (WSNs), contain small sensor nodes, central points for data monitoring. These systems offer various different types of applications for military, transportation, healthcare, agriculture, and industries [1]. The center point is located in the center and has more resources than other sensor nodes in terms of energy, storage, and processing. Due to limited energy resources, in most of the areas,' the replacement of batteries is not possible [2]. Although, sensor-based applications are more useful and smart systems and have more satisfactorily results [3].

In order to fulfill the requirements of the network, there are various unique strategies have reported which are based on sensor networks. Scope control method by using the controlling position of center nodes is one of the significant objectives for WSN for switching the data. Sensor nodes are usually sensing the normal and physical conditions such as sound, temperature, weight, and development or toxic substance. The sensor-based networks are empowering to escalate the spatial scale [4]. The small sensor nodes are communicating with each other and gather the monitored data. Although, these networks have many advantages but still suffered with coverage issues [5]. There is a need to design an algorithm to provide the separation capability among neighbor hubs and fulfilling all imperatives of these networks.

The main objective of this paper is to address the extension issue in WSN. This paper also focuses the Particle Swarm Optimization (PSO) for WSN extension zone.

In addition, this paper illustrates the insurance feature of the algorithm and multiplication of different nodes arrangements including the reason for extension or hindrance scope and full degree.

The rest of the paper is organized as follows:

Section 2 presents the overview of WSN, its applications, coverage area, and recent issues. Section 3 presents the particle swarm optimization. Section 4 presents the deployment of WSN using PSO. Section 5 concludes the paper.

2. Wireless Sensor Network

These networks are self-organized and mostly sensor nodes initiate the decision at self-basis and monitor the circumstances. These networks have faced various challenges including limited resources including memory, interference, and coverage. Particle Swarm Optimization (PSO) refers to an algorithm which offers to streamline and solve the mixture, center point limitation, and batching issues. Authors in [6] access the observed marine microorganism and determine the exactness farming and checking contaminant transport. In business and C-Civil applications for WSN, there is an enthusiasm for structure design for seismologists imagine networks and comprehend the proliferation of quakes at different scales. This engenders the soil conditions and impact on seismic tremors [7]. Due to a large number of sensor nodes, the network cost is at a high peak and nodes cost at low peak. In order to solve this issue, need to find the center points.

2.1. Wireless Sensor Network Issues

To get the optimal coverage is based on the best sensor node distribution in the network. There are some assumption for regular suspicion sensor network application where sensor hubs are able to cover the detecting field or area. The center hubs should cover without any abnormality. In addition, two-dimensional directions of the sensor hubs and their coverage should meet the network area.

2.2. Function Optimization Issue

Execution of an algorithm is based on benchmark function where the variable should be constant for a specific interim. The existing scenario is based on benchmark functions, which incorporates the grievance, Rosenberg, circle and astringent, and mix-streamlining problem. The objective should arrange the space, and improve the mix sorting, screening and order. It is a vital area for operational research.

2.3. Coverage of WSN Network

Sensor nodes are used to detect the nature of any specific event. This area is one of the key areas to offer the Quality of Service (QoS) [8]. Basically, there are three main concerns in WSN including arbitrary sending, constrained easiness go and insufficient sensors to cover the entire ROI. Limited energy resources have a serious impact on network processes and decrease the network scope and not fulfill the converge of the whole ROI. The sensor nodes should be picked and cover a large area is one of the big issues. Another issue is related to the portion of the sensor node due to their separation and irregular organization. In order to address these limitations, this study has adopted the arrangement and scope for a better arrangement of nodes in ROI. Three types of scopes are categorized.

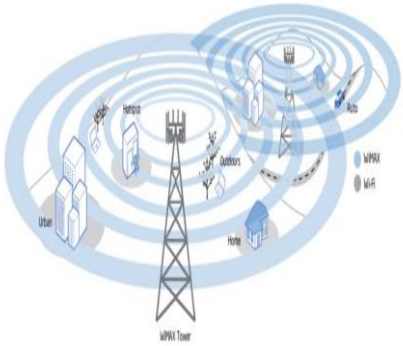


Fig. 1. Area Coverage

In monitored area, the number of pixels denoted with A, where $P_{cov}(C_{ov})$ is used to measure the cover rate of every pixel. The area coverage rate denotes with $c(R_{area}(C))$ is and denoted as a ratio of the covered area of nodes set C, where A denotes the area and m and n defined width and length respectively:

$$R_{area}(c) = \frac{\sum P_{cov}(C_{ov})}{m \times n} \quad (1)$$

2.4. Coverage of Sensor

Sensors coverage issue is one of the considerable issues in WSN which leads to degrades the quality of the service. Especially in complex or critical applications, the coverage is still a challenge. Due to limited resources of sensor nodes and detection certainty in the network, the coverage process tries to search the weak areas and suggest the reconfiguration methods to address the coverage issue. Some authors solved

the coverage issue by adopting or taking the whole coverage area. On the other hand, some have focused on sub-coverage (point and Barrier coverage). Sensor node at the location (X_1, Y_1) covers the (X_2, Y_2) location points if the distance between two points is based on Euclidean distance (Equation 2).

$$(X_1 - X_2)^2 - (Y_1 - Y_2)^2 \leq r^2 \qquad (2)$$

In above Equation, the *r* denotes the sensor nodes radio range where the mean value of the location points (X_i, Y_i) are i= 1,2,3,.....*M*, which presents by (m_x,m_y). Centroid location based nodes cover the distance between most distant area point and sensor hub for detecting the range of the node. Then, the region A which is further categorized into R regions and minimizes the distance between closest centroid and location points. The area *A* covers *R* (sensor nodes). Whereas, the coverage issue denoted by an optimization issue and presents as *P*. The *P* focuses on the arrangement where R is the fixed sensor nodes number. The area is covered as per Equation 3.

$$F = \forall_R \forall_J (\max(distance\,(S_R, P_j)) \qquad (3)$$

The S_R indicates sensor nodes deployment point and P_J is presents the positioning point. The distance has calculates as per Equation 1. The basic objective is to minimizing the F, whereas the sensor sensing range *r* is required to covering the minimum location points.

3. Practical Swarm Optimization

Pragmatic Swarm Optimization (PSO) is another approach designed by Kennedy in [9]. In 1995, the swarm has adopted as one of the optimal approach [10]. They used this strategy and defined five basic standards and for fake life application for swarm knowledge. Basically, the molecule was alluded with a speed to increase the velocities and volume and mass-less. Molecule swarm enhancement is one of the powerful compared to other algorithms.

3.1. Method:

The assumption is based on 20x20sm (square meter) area. This area has number of pixels where every pixel is 1-sm. Figure 2 presents the nodes positions.

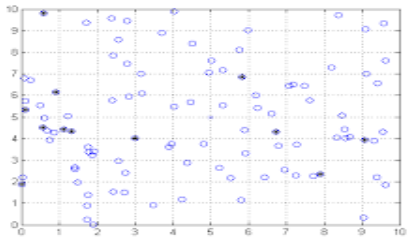


Fig. 2. 20 sensor nodes deployed randomly.

Probability is adopted for this model where coverage area is taking into account. For coverage area rate, following points are considered:

1. Calculation of coverage area by determine coverage rate of one pixel for every sensor node.
2. Calculation of joint coverage rate of every pixel.
3. For calculate the joint coverage rate of every pixel need to repeat step 1 and 2.

In order to calculating the coverage area by using coverage control algorithm. There is an assumption where one group has M particles and every particles is based on D dimensional units with different positions, for I particle, Xi shows its position, *Xi = (X1, Xtwo, Xthree, ...Xd)* with speed of *Vi = (V1, V2, V3, ...V4)* respectively. The optimal position P_{Best} which the unit passes is $P_i = (P_{i1}, P_{i2}, P_{i3}, \ldots\ldots\ldots, P_{id})$ and P_g denoted best position.

3.2. PSO Algorithm and General Code

The PSO principle is that where every node independently seen as particle. The particle has no volume multidimensional search space. This particle is flying at certain speed. The speed is always dynamic based on its flying experiences and group experience, this defines as:

1. $X_i = (X_{i1}, X_{i2}, \ldots\ldots X_{in})$ is the *i-th* particle.
2. $V_i = (V_{i1}, V_{i2}, \ldots\ldots, V_n)$ current flying speed of the *i-th* particle.
3. $P_i = (P_{i1}, P_{i2}, \ldots\ldots\ldots, P_{in})$ is the best position in the of *i=th* particle, where place has best value for the experience of the *i-th* particle.

3.3. PSO algorithm Principles:

There are five articulated discussed in [11], which are as follows:

1- Proximity: the population of nodes offer computations.
2- Quality: quality factors are taking into account for adjustment in the network.

3. Diverse response: the nodes population is not confine itself into narrow channels.
4. Security: the continuous change of nodes and population always tackle the security concerns.
5. Adaptability: nodes population is always providing alter behavior where the energy consumption and resource is worthy.

4. Deployment of WSN using PSO

The sensor deployment issue is one of the challenges where the nodes determine the positions for its sufficient operations. This process costumes energy and computation resources of nodes [9]. The problem is that most of the sensor nodes remain unnoticed about events in the network. The complex networks with more nodes have suffered from congestion and delay issues. The best deployment of sensor nodes ensures the services quality, network lifetime and cost. Available PSO solutions are centrally deployed base station to determine the positions of sensor nodes, mobile nodes or base stations.

4.1. Coverage Problem using PSO

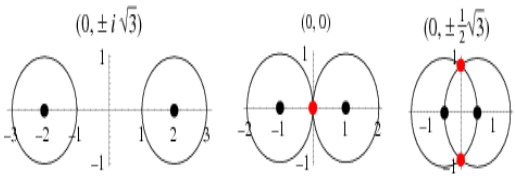


Fig. 3. Circle-circle intersection

The two circles are intersected into two imaginaries, which are single or distinct point. The intersection line between two circles known as radical line, whereas the three circles intersected in single point, intersection of circle pairwise radical lines called radical center.

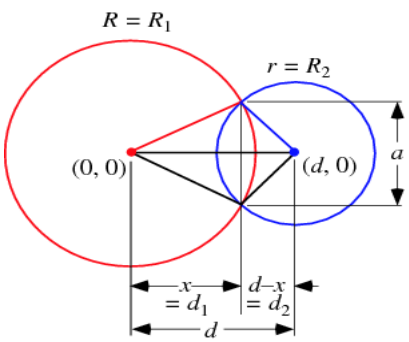


Fig. 4. Radical Center

The two circles of radii R and r centered at (0, 0) and (d, 0) intersect in region shaped like an asymmetric Figure 4 (Equation 4,5).

$$x^2 + x^2 = R^2 \quad (4) \qquad (3)$$

$$(x-d)^2 + y^2 = r^2 \quad (5) \qquad (4)$$

After merging 1 and 2:

$$(x-d)^2 + (R^2 - x^2) = r^2 \quad (6) \qquad (7)$$

Multiplying through and rearranging gives:

$$x^2 - 2dx + d^2 - x^2 = r^2 - R^2 \quad (7)$$

Solving for X results in:

$$x = \frac{d^2 - r^2 + R^2}{2d} x = \frac{d^2 - r^2 + R^2}{2d} \qquad (8)$$

Discussion

The simulation results indicated that the algorithm successfully found the best deployment strategy for this application. The position determine through sensor node range to fulfill the goal of maximum coverage, rd connecting the cusps of the lens therefore has half-length *y* given by plugging *x* back in to obtain.

$$y^2 = R^2 - x^2 = R^2 - \left(\frac{d^2 - r^2 + R^2}{2d}\right)^2 \quad (9)$$

$$= \frac{4d^2R^2 - (d^2 - r^2 + R^2)^2}{4d^2} \quad (10)$$

Solving for x and plugging back in to given entire chord length a=2y then gives:

$$\frac{a = 1}{d}\sqrt{4d^2R^2 - (d^2 - r^2 + R^2)^2} \quad (10) \qquad (11)$$

$$= \frac{1}{2}\sqrt{(-d+r-R)(-d-r+R)(-d+r+R)(d+r+R)} \qquad (12)$$

This same formulation applies directly to the sphere-sphere intersection problem.

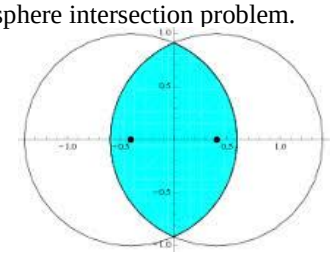


Fig. 5. Area of Two Unit Disks.

In order to find optimal coverage of particles and coverage of nodes and their radius denoted with (Rs) are equal with the following steps:

1- The sensor node are distributed randomly.
2- Whenever, the distance between two sensor nodes are greater than (2R) then the intersection exists.
3- Determine the every node distance with its neighbor node and whenever the distance of two nodes is equal to R2 then no intersection exists.
4- Whenever, the distance between nodes is less than (2R), then the distance towards intersection consider large.
5- To adopt same calculation for other sensor nodes, we determine the nodes instruction and multiplying every time based on changings the nodes distribution till find the optimal coverage using PSO approach. Figure 7 presents the coverage of nodes and their distance calculation.

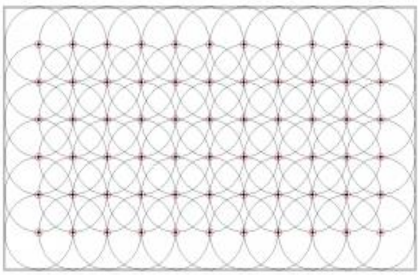


Fig. 6. Best coverage sensors

N= nodes
R=radius of sensor
Rs=sensing radius
Area=M*M
N=4*Area/2πr^

$$d^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 \qquad (12)$$

When d >2R (no intersection between two nodes).
When d < 2r (intersection).
When d = 2r (no intersection).

Case Study

Many researchers have explored the mirror based on swarm streamlining for WSN. An approach is presented based on PSO and wellness function in [10]. This approach limits the introduction bunch and removes it among the group and nodes hubs. This approach advances the vitality productivity. On the opposite side, PSO also executed at BS level. At the initial stage (every bunch head round), each sensor node sends its own zone information and remaining essentialness level to the SN (sink node). This type of transmission expands clog and depleting the pointless vitality. In any situation, the reenactment initiates for the technique for PSO-grouping or PSO-C.

Discussion

The simulation results indicated that the algorithm found the optimal approach for this application where the sensor hub position detected and able to provide the most extreme scope as presented in Figure 6 and Table 1. The scope rate increments as the detecting span *r* increments, while the cycles diminish as the detecting sweep r increments, when r compasses to 5m, then target the area which is completely covered, according to the results of the simulation, the experimental data table can be made as below:

Table 1. Results of sensing radius, iteration and coverage rate

Sensing Radius	1.5	2	2.5	3	4	5
Iteration	130	100	76	61	55	48
Coverage Rate	22.93%	39.98%	58.68%	76.53%	94.13%	100%

5. Conclusions

In this paper, the centralized deployment algorithms are discussed which are basically work as a powerful node. The powerful node means has more computation and communication capabilities compared to other nodes in the network. With many advantages, implemented centralized algorithms in these nodes have suffered from scalability issues. Scalability issue is increasing more especially for large size networks where various sensor nodes are working together. Centralized node has more responsibilities due to the large size of messages handling and becomes a bottleneck. In order to address the centralized node issues, the distributed algorithms have been designed which are usually based on processing, routing decisions, and moving capabilities. However, these features might not be feasible. The distributed algorithms are expensive in terms of large data handling and energy consumption, especially in highly dynamic networks. In the future, we will focus to design more appropriate approach which will be based on intersections areas and able to select control parameters for PSO.

References

[1] K. N. Qureshi, A. H. Abdullah, and R. W. Anwar, "Wireless sensor based hybrid architecture for vehicular ad hoc networks," TELKOMNIKA (Telecommunication Computing Electronics and Control), vol. 12, pp. 942-949, 2014.

[2] E. Fadel, V. C. Gungor, L. Nassef, N. Akkari, M. A. Malik, S. Almasri, et al., "A survey on wireless sensor networks for smart grid," Computer Communications, vol. 71, pp. 22-33, 2015.

[3] D. Bell and M. A. Leabman, "System and method for smart registration of wireless power receivers in a wireless power network," ed: Google Patents, 2017.

[4] M. Anwar, A. H. Abdullah, K. N. Qureshi, and A. H. Majid, "Wireless Body Area Networks for Healthcare Applications: An Overview," Telkomnika, vol. 15, pp. 1088-1095, 2017.

[5] S. Iqbal, A. H. Abdullah, and K. N. Qureshi, "Channel quality and utilization metric for interference estimation in Wireless Mesh Networks," Computers & Electrical Engineering, vol. 64, pp. 420-435, 2017.

[6] M. F. Othman and K. Shazali, "Wireless sensor network applications: A study in environment monitoring system," Procedia Engineering, vol. 41, pp. 1204-1210, 2012.

[7] J. Caffrey, R. Govindan, E. Johnson, B. Krishnamachari, S. Masri, G. Sukhatme, et al., "Networked sensing for structural health monitoring," in Proceedings of the 4th International Workshop on Structural Control, 2004, pp. 57-66.

[8] H. Zhang and J. C. Hou, "Maintaining sensing coverage and connectivity in large sensor networks," Ad Hoc & Sensor Wireless Networks, vol. 1, pp. 89-124, 2005.

[9] J. Kennedy, "Particle swarm optimization," in Encyclopedia of machine learning, ed: Springer, 2011, pp. 760-766.

[10] M. M. Millonas, "Swarms, phase transitions, and collective intelligence," arXiv preprint adap-org/9306002, 1993.

[11] J. Kennedy and R. Eberhart, "Particle swarm optimization inProceedings of IEEE International Conference on Neural Networks," Piscataway December, 1995.

A Survey of Deep Learning Techniques used for Ransomware Attacks Detection

[1]Umara Urooj, [1]Mohd Aizaini Bin Maarof, [2]Bander Ali Saleh Al-rimy
[1]School of Computing, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia
[2] Faculty of Business and Technology, Unitar International University, 47301 Selangor, Malaysia
[e-mail: umaraurooj@gmail.com, aizaini@utm.my, bnder321@gmail.com]
*Corresponding author: Umara Urooj

Abstract

Ransomware became prominent field for the researchers in the past few years due to rising number of ransomware attacks. Ransomware attacks are increasing due to its easy to create mechanisms. The irreversible nature of ransomware attacks making it more dangerous for the user files. Deep learning concepts are now extensively used in the field of ransomware attacks detection due to its capability of optimal feature extraction from raw inputs. Initially, this paper has a brief description of concepts involved deep learning studies. In the next section a survey about deep learning solutions used in the field of ransomware attack detection is presented.

***Keywords*:** Security, Malware, Ransomware, Detection, Deep learning, Attacks

1. Introduction

Security is major concern in all the ages. Protecting data from the malware attacks defined a separate domain in the field of computing. Malware is a software which damages the user computational resources, steals or damage data or sometimes make it inaccessable. There are many types of malware include Trojans, Viruses, Ransomware, Spyware and more [1, 2].

The term Ransomware is defined as a malicious software that demands ransom in exchange of access to your legitimate computing resources i.e. user files [3]. It usually demands ransom by displaying a message on the screen and making you unable to access any other thing on your device. This is one of the most prominent malware now a days due to moneymaking game. Its easy to install mechanism, making it easy to spread on multiple victims machines. It can easily spread through deceptive links, emails, websites and intant messages. Once a device exposed to ransomware attack there's no other way to access files again except decrypting it from decryption key [4].

According to mcafee report Ransomware attacks are growing at fast rate. New ransomware families are detected with an increase of 118% ransomware attacks due to development of innovative techniques. Ransomware grew by 118% in the first three months of 2019. This malware is spreading due to easy to launch mechnasims [5].

Internet is one of the most needed thing of new age. The prominent use of internet made the users more vulnerable. With a little knowledge a ransomware attack can be launch. Due to increasing number of attacks the ransomware maintains the attaention in the field of cyber secuirty. Benign like nature of ransomware attacks makaing it more challenging to deal with. Innovative methods to spread the ransomware attack are increasing due to the money incentive attached to the cause damage.

Machine learning and deep learning are being widely used in detection of malicious softwares i.e.

malwares and ransomware due to their effective techniques. New researches are conducted to get more benefits of these techniques to overcome the ransomware attacks [6]. The existing models have limitations in terms of dealing with high dimesnsional data and high false alarms. There is still need to design the more efficient detection model which can confront the gradually developing rasomware attacks while considering the limitations of existing solutions.

The key contribution of this paper includes:

- Presentation of prominent Deep Learning solutions available for ransomware detection.
- Summary table of DL methods.
- Comparision on the basis of different parameters including types of platform, types of analysis and DL methods used.
- Applications of DL solutions for different types of ransomware attacks.

The paper is organized as section 2 described related concepts; section 3 briefly explained the available studies related to ransomware attacks detection by utilizing deep learning concepts. Section 4 discussed available solutions for ransomware detection. Conclusion of the paper is discussed in the end in section 5.

2. Concepts and Descriptions

2.1 Ransomware Analysis

Ransomware analysis plays a vital role in the detection of ransomware attacks. Ransomware analysis helps to figure out the behavior of a malicious program and type of operations associated with the malicious program [7].

2.1.1 Static Analysis

Static analysis is a passive approach perform by executing the malicious code. It works by matching or comparing the features of malicious code with already available features [7, 8].

2.1.2 Dynamic Analysis

Dynamic analysis is an active approach performed by executing the malicious code. Malicious code run under the controlled environment and features exhibited are captured in the controlled environment [7, 8].

2.1.3 Hybrid Analysis

Hybrid analysis is an analysis approach which take advantage of static and dynamic analysis to produce more accurate analysis [9, 10].

2.2 Ransomware Detection

Ransomware detection is aimed at finding out the ransomware attack during or after it takes place. There are two approaches for ransomware detection i.e. structural and behavioural [7].

3. Related work

Serval researchers were conducted as survey and review for DL technologies. Following discussed studies will highlight the difference of presented survey from the existing literature.

In [11] a review of 8 papers for android platform using deep learning for ransomware detection was presented. The study highlighted the deep learning techniques used for android platform.

A review was presented about machine and deep learning techniques for the IoT devices. Machine and deep learning methods were discussed in detail with respect to security on IoT layers [6].

The study of [12] reviewed the deep learning threats, attacks and deffensive techniques. The study divided the secutiy attacks into two categories from which the attacks related to deep learning were outlined.

In [13] the authors presented a survey for deep learning methods for cyber security applications. The study discussd the datasets and metrics used for deep learning evaluation.

4. Deep Learning methods for ransomware detection

Serval studies were conducted which utilized the deep learning concepts in the detection of ransomware attacks. Some of the available detection models are discussed below:

In [14] a ransomware detection model was proposed named Dna-Droid. System used deep learning to label the data into benign and

ransomware category. Two types of analysis were carried out to analyze samples. First static analysis is performed then samples were fed to dynamic unit for classification. This model took data in the form of API calls, text and images. Sequence of API calls was observed to monitor samples. API calls were compared to the calls that malicious program generated.

The study of [15] a deep neural network based ransomware detection model was introduced. It was a network based detection system that observed suspicious process by network calls from malicious application. Model was trained by using the genrated HTTP requests from malicious application. This study achieved 93.9% accuracy rate by correctly identifying the ransomwares.

A deep learning based method Long Short Term Memory (LSTM) was utilized in [16] for the detection of ransomware attacks. The system dynamically analyzed the samples. API calls were observed for accurate detection of ransomware attacks.

In [17], a deep learnng based ransomware detection model was presented. It utilized LSTM with attention mechanism to capture the patterns of ransomware behaviour. This model provided the better results then the existing available solutions.

The authors of [18] proposed a ransomware detection model using CNN, LSTM and OCSVM. LSTM provided the best performance for vectoring the activites of suspecious applications. Generated vector was fed to the OCSVM for ransomware family identification. A combination of LSTM and CNN was used for classifying the malicious program. This system has limitation in terms of its complex processing.

The study of [19] proposed a static analysis framework for the detection of fingerprinting ransomwares. N-gram opcodes sequences were statistically analysed by using deep learning and self-attention mechanisms for the detection of fingerprinting ransomwares. This framework has limitation of hardly handling the advance packing techniques due to the use of static analysis.

RanSD (Ransomware static and dynamic analysis) technique was proposed in [20]. It utilized static and dynamic analysis for the detection of ransomware by performing feature engineering. This technique focused on finding the most discriminative features to increase the accuracy. N-Gram technique was utilized to find the most important sequence. Deep convolutional neural networks and conventional machine learning technique was used to train the ransomware detection model.

A detcetuon framework RAPPER was presented in [21]. This study presented a two-step preventive detection tool RAPPER for the detection of malicious process in less than 5 seconds. System provided backup of files if a ransomware attack was encountered. It utilized LSTM based autoencoder with minimal trace points. The system worked in both online and offline states. It was a desktop based ransomware detection technique which distinguish the disk encrypting programs and a malicious program. The processes performance was observed in terms of statistics by using hardware performance counters. The autoencoders developed the profile for normal process which helped to detect the ransomwares.

In [22] a ransomware detection method was proposed based on memory storage. This method was designed for the devices with limited computation resources like IoT and embedded systems. Model was built by utilizing the stochastic computation and deep belief network structure. Model improved accuracy and speed by achieving 91% precision and 0.006ms detection speed.
Memory-Assisted-Stochastic-Dynamic-Fixed-Point arithmetic method produced cross-correlation for the stochastic computation in FPGA.

In the study [23] a deep learning-based detection model was developed DeepRan for ransomware early detection in operational enterprise system. High dimensional Host logs from bare metal servers were used to extract the features for training of detection model. The host profiles are designed by attain based BiLSTM Recurrent Neural Network. Study results were compared with other five deep learning models.

4.1 Analysis of Deep Learning techniques

This section summerized the studies considered in this paper. Deep learning techniques used for the detection of ransomware attacks along with other considered parameters are breifly described in the table below.

Table 1.1 Analysis of Deep Learning techniques

Approaches	Platforms	DL Methods	Analysis Type	Evaluation metrics
Dna-Droid [14]	Mobile & Windows	Deep Autoencoder	Hybrid	Accuracy, Precision, Recall, FP
Not mentioned [15]	Network	DNN	Dynamic	Accuracy TN, FP
Not mentioned [16]	Desktop	LSTM	Dynamic	Accuracy
Not mentioned [17]	Desktop	LSTM	Static	Accuracy, FP
DRTHIS [18]	Fog layer	LSTM DNN	Static	F-measure, Precision, Recall, TP, TN, FP, FN
Not mentioned [19]	Desktop	SA-CNN	Static	Accuracy, Recall, Precision, F1-score
Not mentioned [20]	Desktop	Transfer Learning based Deep Convolutional Neural Network	Hybrid	Accuracy, Recall, Precision
RAPPER [21]	Desktop	LSTM	Dynamic	Detection Speed
Not mentioned [22]	Embedded system & IoT	DBN	Dynamic	Accuracy, Detection speed
DeepRan [23]	Bare Metal Server OS	LSTM	Dynamic	Accuracy, F1 score, Precision, Recall, MCC

As shown in **Table 1.1**, different platform used different deep learning methods. Most of the detection system utilizing deep learning were developed for desktop computers. Deep learning method LSTM was utilized more due to its efficient work and control over entering input. It stores information for long time. Most of the studies used dynamic analysis for ransomware detection. The detection models are mostly evaluated in terms of accuracy, precision, recall, TP, TN, FP, FN.

5. Conclusions

In this paper, we presented a survey about the deep learning concepts utilized in the detection of ransomware attacks. In this survey, different types of ransomware attacks and available deep learning solutions to counter those attacks, are discussed. Each ransomware detection solution is compared in the end. This survey aims to provide a useful manual that can encourage researchers as a direction to work with deep learning technologies in the field of ransomware attack detection. It can help them in developing the more efficient detection model.

Appendix

DBN: Deep Belief Network
DNN: Deep Neural Network
CNN: Convolutional Neural Network
SA-CNN: Self-Attention based Convolutional Neural Network
TP: True Positive
TN: True Negative
FP: False Positive
FN: False Negative
OS: Operating System
SVM: Support Vector Machine
LSTM: Long Short Term Memory
MCC: Matthews Correlation Coefficient
FPGA: Field Programmable Gate Arrays
OCSVM: One-Class Support Vector Machine

References

[1] A. Software, "What Is Malware? How Malware Works & How to Remove It ".

[2] B. A. S. Al-rimy *et al.*, "Redundancy Coefficient Gradual Up-weighting-based Mutual Information Feature Selection technique for Crypto-ransomware early detection," *Future Generation Computer Systems,* vol. 115, pp. 641-658, 2021/02/01/ 2021.

[3] Y. A. Ahmed, B. Koçer, S. Huda, B. A. Saleh Al-rimy, and M. M. Hassan, "A system call refinement-based enhanced Minimum Redundancy Maximum Relevance method for ransomware early detection," *Journal of Network and Computer Applications,* vol. 167, p. 102753, 2020/10/01/ 2020.

[4] K. Lab, "What is Ransomware?."

[5] C. Beek *et al.*, "Mcafee labs threats report: August 2019," *McAfee Labs,* 2019.

[6] M. A. Al-Garadi, A. Mohamed, A. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for internet of things (IoT) security," *IEEE Communications Surveys & Tutorials,* 2020.

[7] B. A. S. Al-rimy, M. A. Maarof, and S. Z. M. Shaid, "Ransomware threat success factors, taxonomy, and countermeasures: A survey and research directions," *Computers & Security,* vol. 74, pp. 144-166, 2018.

[8] S. Kok, A. Abdullah, N. Jhanjhi, and M. Supramaniam, "Ransomware, threat and detection techniques: A review," *Int. J. Comput. Sci. Netw. Secur,* vol. 19, no. 2, pp. 136-146, 2019.

[9] M. Choudhary and B. Kishore, "HAAMD: hybrid analysis for Android malware detection," in *2018 International Conference on Computer Communication and Informatics (ICCCI)*, 2018, pp. 1-4: IEEE.

[10] A. Alqahtani, M. Gazzan, and F. T. Sheldon, "A proposed Crypto-Ransomware Early Detection(CRED) Model using an Integrated Deep Learning and Vector Space Model Approach," in *2020 10th Annual Computing and Communication Workshop and Conference (CCWC)*, 2020, pp. 0275-0279.

[11] N. Alzahrani and D. Alghazzawi, "A Review on Android Ransomware Detection Using Deep Learning Techniques," in *Proceedings of the 11th International Conference on Management of Digital EcoSystems*, 2019, pp. 330-335.

[12] M. I. Tariq *et al.*, "A Review of Deep Learning Security and Privacy Defensive Techniques," *Mobile Information Systems,* vol. 2020, 2020.

[13] D. S. Berman, A. L. Buczak, J. S. Chavis, and C. L. Corbett, "A survey of deep learning methods for cyber security," *Information,* vol. 10, no. 4, p. 122, 2019.

[14] A. Gharib and A. Ghorbani, "Dna-droid: A real-time android ransomware detection framework," in *International Conference on Network and System Security*, 2017, pp. 184-198: Springer.

[15] A. Tseng, Y. Chen, Y. Kao, and T. Lin, "Deep learning for ransomware detection," *IEICE Tech. Rep.,* vol. 116, no. 282, pp. 87-92, 2016.

[16] S. Maniath, A. Ashok, P. Poornachandran, V. Sujadevi, A. P. Sankar, and S. Jan, "Deep learning LSTM based ransomware detection," in *2017 Recent Developments in Control, Automation & Power Engineering (RDCAPE)*, 2017, pp. 442-446: IEEE.

[17] R. Agrawal, J. W. Stokes, K. Selvaraj, and M. Marinescu, "Attention in Recurrent Neural Networks for Ransomware Detection," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 3222-3226: IEEE.

[18] S. Homayoun *et al.*, "DRTHIS: Deep ransomware threat hunting and intelligence system at the fog layer," *Future Generation Computer Systems,* vol. 90, pp. 94-104, 2019.

[19] B. Zhang, W. Xiao, X. Xiao, A. K. Sangaiah, W. Zhang, and J. Zhang, "Ransomware classification using patch-based CNN and self-attention network on embedded N-grams of opcodes," *Future Generation Computer Systems,* vol. 110, pp. 708-720, 2020.

[20] A. Ashraf, A. Aziz, U. Zahoora, M. Rajarajan, and A. Khan, "Ransomware Analysis using Feature Engineering and Deep Neural Networks," *arXiv preprint arXiv:1910.00286,* 2019.

[21] M. Alam, S. Sinha, S. Bhattacharya, S. Dutta, D. Mukhopadhyay, and A. Chattopadhyay, "RAPPER: Ransomware prevention via performance counters," *arXiv preprint arXiv:2004.01712,* 2020.

[22] K. Alrawashdeh and C. Purdy, "Ransomware detection using limited precision deep learning structure in fpga," in *NAECON 2018-IEEE National Aerospace and Electronics Conference*, 2018, pp. 152-157: IEEE.

[23] K. C. Roy and Q. Chen, "DeepRan: Attention-based BiLSTM and CRF for Ransomware Early Detection and Classifcation," *Information Systems Frontiers,* pp. 1-17, 2020.

Recent Advancement in Wireless Sensor Networks: Model, Architecture, Standards, Challenges and Future Aspect

Hawezi Omeed kamal khorsheed, MD Asri Bin Ngadi, Abdul Hanan Abdullah and Israr Ghani
School of Computing, Faculty of Engeering, Universiti Teknologi Malaysia
Skudai, Johor Bahru, 81310, Malaysia
[Omeed.kamal@graduate.utm.my, hanan@utm.my, dr.asri@utm.my, israrghani@outlook.com]
*Corresponding author: omeed kamal

Abstract

Wireless Sensor Network (WSN) is made up of a group of autonomous mobile nodes which are communicating without the requirement of existing infrastructure. The major problems observed in WSNs are bandwidth limitations, battery power and frequently changing topology. To handle these problems in WSNs multicasting is required preferably. Multicast routing is capable of maintaining link in dynamic changing topology. It is very difficult and challenging task to achieve energy efficient routing in WSNs. It is inadequate for WSNs to communicate long time with limited battery power. To get optimized operation of wireless nodes data should be routed in such a way that energy consumption can be minimized. To achieve these objective many energy efficient multicast routing protocols has been proposed. This paper provides a review of various WSNs multicast routing protocols that are focusing on energy efficiency. Also Proposed a solution to achieve energy efficiency after comparative study analysis

***Keywords*:** Wireless Sensor Network, Multicast Routing, Energy Efficiency, Bandwidth, Optimized Routing

1. Introduction

A wireless sensor network is consisting of a group of independent mobile nodes which are communicating wirelessly without the requirement of existing infrastructure. WSNs are increasingly popular in recent years due to their impending use in emergency environments such as battle fields, disasters etc. Mostly WSNs are designed for scalar data such as temperature, pressure, humidity and location etc.

To minimize the energy consumption is a major factor to be considered in WSNs. Most of the sensor networks uses multicasting for routing between source and sink because multicasting utilizes the resources economically. Limited bandwidth, battery power and frequently changing topology of the nodes are those concerns that make the multicast routing challenging. Multicast routing is used between active groups in which a single source can contact to multiple receivers but it is a very difficult task to achieve a reliable link and utilize battery power efficiently. Limited battery power results in limited node life which onward reduces network operational time. To efficiently utilize the battery life of a node, low energy consumption protocols need to be proposed and adopted.

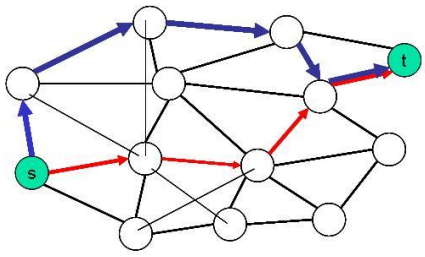


Fig. 1. Wireless Sensor Network

Multicast routing is the one of best techniques to achieve this objective as it use several paths between source and destination which are intended to provide reliability, high throughput and low energy consumption by identifying best routes to the destination. WSNs can be used in current scrutiny systems. It can improve the ability of LEAs by monitoring public events, border areas and public installation. WSN can be helpful in different activities such as traffic violations, car parking etc. [1].

This paper provides a wide comparison of various protocols that deals in WSNs focusing on energy efficient multicast routing. In addition, paper also discuss the WSN model, architecture, standards and challenges.

The rest of the paper is organized as follows. Section 2 present related works through surveys of multicast routing and energy efficient protocols. In section 3 some problem is described which are observed in the reviewed papers. Section 4 provides a comparative analysis of different protocols. Section 5 describes some techniques and procedures for energy efficiency. In section 6 a proposed solution is introduced. Section 8 concludes the paper with future direction.

2. Related Work

Multicast improves the efficiency of the wireless link as whole by exploiting the broadcast property of wireless transmission. In recent years many multicast protocols have been proposed for WSNs. These protocols have different features and mechanism. Here a survey of some distinctive multicast routing protocol has been provided, with view of how to get ride on the limitations presented in these proposed protocols.

2.1 On-Demand Multicast routing protocol (ODMRP)

S. J Lee et al proposed ODMRP [2] protocol which is a source-based multicast routing protocol. It describes the idea of a forwarding group. When a source has data to transmit and it doesn't have any routing path, it will broadcast a JOIN data packet. On receiving a join request the receiver stores the upstream address in its routing table. Update the join request and rebroadcast the request packet. Any node that want to join the group, broadcast a JOIN reply to the request. When a node receives a join reply, it set its forwarding flag and rebroadcast after updating the routing table of next node. Join requests are refreshed periodically by source and reply to this request is must for session members to remain in group. If the reply to periodic refresh packet is not received, then the node is repelled from forwarding group nodes and forwarding group reset their flag according to updated state.

2.2 Enhanced-ODMRP

J. S Park et al proposed E-ODMRP [3] which is an enhancement of ODMRP protocol. This technique alter the refreshing period dynamically instead of mobility prediction which is useful to minimize the flooding. And also it will avoid link breakage or at least minimize this activity. Through this way E-ODMRP enhance the performance of this protocol. Overhead through this protocol is reduced up to 50% and packet delivery is same as in the ODMRP.

2.3 Core-Assisted Mesh Protocol (CAMP)

J. J Garcia et al proposed this protocol [4] which is an extension of the core based tree (CBT). To keep all the multicast groups connected, a shared core has been used in this protocol. This core has the ability to connect in fast mobility of the nodes. Multiple cores are established for multicast groups for JOIN operation instead of flooding. It has receiver based JOIN procedure to join any suitable multicast group. If a node wants to join a multicast group, it will transmit a JREQ to the core and nearest multicast group send a RREP in response. If the node is a member of any existing multicast group, then it announces its affiliation. The node broadcast JREQ if the core is out of reach. The big flaw in this protocol is continuous refresh massaging to maintain the core which results in more energy consumption and hence minimizes the network life.

2.4 Associativity-Base Ad-hoc multicast protocol (ABAM).

S. Bunchua et al proposed ABAM [5] protocol. It builds a source-based multicast tree. When many beacons are received uninterruptedly from the neighbor nodes, the source reaches a

threshold level and the source node gains stable association of neighboring nodes through this process. Which helps the source to select route which will be more likely has longer life and will require less configuration. To join a multicast tree, a request is forwarded, and as a result replies are received from group members. The requester selects a best rout and sends an approval. The source initiated the formation of tree and selects a rout with highest association at ability. When the association is violated, there occurs tree reconfiguration. To leave the tree, an announcement is broadcasted upstream along the tree until the receiving node is reached.

2.5 Power-aware localized routing (PLR).

Xu Lin et al proposed PLR [6] protocol which is a fully distributed energy aware algorithm and restricted locally in sense of routing. The information of neighbors and destination are assumed to be known by the source. Corresponding link cost from source to its destination through its neighbors is known. Depending on this information the optimized path cannot be achieved, but the source picks the next hop through this process over all transmission power of the sender node can be minimized. It save energy but has poor throughput and more delay.

2.6 Lifetime-aware Multicast Tree (LMT).

LMT [7] finds route that minimize the difference of residual energy of the node in the network and in result maximize the overall life of the network. It assumes that the transmission energy is directly proportional to the distance from the sender to the receiver. In comparison with existing multicast algorithms, the LMT proves its efficiency.

2.7 Power-aware Multiple Access protocol (PAMAS).

M. Maleki et al proposed PAMAS [8] protocol. It uses routing algorithm that saves the energy of the nodes that are not in active role like AODV but it not utilizes the low energy nodes for execution. Through this technique it improves the life of a network significantly. PAMAS has a minor drawback of packet delivery and delay at low traffic rate. But at high rate both these factors are improved due to reduced congestion. PAMAS follows the AODV but it can be implemented on any on demand protocol.

2.8 Probabilistic Predictive Multicast Algorithm (PPMA)

D. Pompili et al proposed [9] this protocol for multicasting. PPMA maximizes network life time by estimating statistically the future position of the nodes. This algorithm has two versions centralized and distributed. It is a tree based algorithm which defines the route on the cost factor based on residual energy, link availability and mobility factor. In centralized approach for a given no. of hops up to the source, the receiver has a set of father nodes to communicate in the tree. In distributed approach, to find a multicast path, a private cost is defined along with the public cost of the paths. Through this approach a new receiver node searches suitable public cost path and join the tree. Old receiver node alter its path when it finds that the private cost is lower than adopted public cost. Flaws in the PPMA are unreliability and lack of robustness. It also keeps all nodes active for continuous tracking to find suitable probabilistic link which consume more energy.

2.9 On-Demand Global Hosts for Mobile Ad-Hoc Multi Cast (OGHAM) Protocol.

G.H. Chen et al proposed this mechanism [10] which has a two tier architecture. When multicasting is required it selects a backbone host for this purpose. A host that has less no of hops to other hosts is selected as Base Host (BH) to get shorter route. BHs have the responsibility to determine the suitable route, packet delivery, group membership and route updating. In the initialization phase the source ‘S’ try to find a BH within a first tier that is a distance of radius of 2 hops, if find then attach itself with this BH. If not found then it broadcast a join massage in broad region, which is second tier greater than 2-hop distance. In response the BHs send their neighboring information to ‘S’. On the basis of this information ‘S’ attach to a suitable BH. OGHAM reduces the transmission time and packet loss but has high overhead and less network life time.

3. Problems in Existing Protocols.

In past few years' different multicast protocols have been proposed for WSNs. There occur so many issues in the proposed solutions which demands power management and QOS in wireless sensor networks. Some notable problems in these reviewed protocols are highlighted below.

a) The ODMRP requires periodic control packet flooding to build a multicast route. Through these flooding it repairs the link break.
b) The ABAM also use periodic neighbor sensing frequently to repair and detect the link break.
c) The big flaw in CAMP protocol is continuous refresh massaging to maintain the core which results in more energy consumption and hence minimizes the network life.
d) LMT enhance the multicast life time through searching optimized routing path that minimize the difference of residual level of energy in the nodes of network.
e) OGHAM has high overhead and less network life time.
f) Flaws in the PPMA are unreliability and lack of robustness. It also keeps 0all nodes active for continuous tracking to find suitable probabilistic link which consume much energy.
g) PAMAS has a minor drawback of packet delivery and delay at low traffic rate.
h) PLR save energy but has poor throughput and more delay.

In a sense it can be said that all of the above protocols are not able to handle the problems lonely like bandwidth, topology and energy consumption.

4. Comparative Analysis

In this section we present an analysis of different wireless sensors protocols proposed by different researchers. Energy efficiency, latency, throughput and various other parameters are taken into account to evaluate and compare the performance of these protocols. The performance of different multicast protocols has been evaluated in accordance with some parameters given in the table below.

Table 1: Protocols in WSNs

S. No	Protocol	Throughput	Delay	Latency	Consumption	Overhead
1	ODMRP	High	Moderate	Low	Moderate	High
2	E-ODMRP	High	Moderate	Low	Low	Low
3	CAMP	Low	Moderate	Moderate	Moderate	High
4	ABAM	High	Moderate	Moderate	High	High
5	PLR	Moderate	High	High	Low	Low
6	LMT	Moderate	High	High	Low	High
7	PAMAS	Moderate	Moderate	Moderate	Low	Low
8	OGHAM	High	Moderate	High	Low	High
9	PPMA	Low	High	High	Moderate	High

5. Suggestions to Reduce Energy Consumption

After review of many proposed energy efficient protocols, some techniques are presented here which are helpful to achieve energy efficiency in WSNs. These techniques can be utilized to achieve energy efficiency in multicast routing protocols, which are used to deliver the data contents in WSNs.

5.1 Avoid useless activity: -

It's not necessary to use high power operational mode when operation can run on lower power mode. Also there is no need to put error control on error-resilient data.

5.2 Scheduled communication: -

communication should be scheduled at suitable time such that the residual power of the nodes can be exploited as much as possible.

5.3 Reduced amount of data: -

This technique provides the compromise between communication and computation. Energy spend by wireless medium is due to communication and due to control information or processing of data is spent by computation.

5.4 Power Management: -

Through this technique periods of idleness are exploited. It activates sleep mode when the system is not in active use.

6. Proposed WSN Model

Keeping in the view the problems and efficiencies of the existing protocols, a proposed solution is presented in this paper. Location-aware Adoptive Multicast Routing Protocol (LAMRP) is the proposed protocol which focuses on route stability and energy efficiency. LAMRP consists of three types of nodes

- Sensor Nodes
- Carrier Nodes
- Sink Nodes

6.1. Sensor Nodes

These nodes sensed data from environment in which they are existing and multicast an awake massage to the carrier nodes. This massage initiated automatically on retrieval of data and contains data type and sink location. On the basis of carrier reply, the sensor node decides a suitable route for data delivery.

6.2 Carrier Nodes

These remains in sleep mode. When a sensor node retrieve data from environment, carrier nodes send a reply. The reply contains energy level and number of hops to destination along with ACK.

6.3. Sink Nodes

LAMRP contains multiple static sink nodes up to 5, deployed at equal distance from each other. They are interconnected, if anyone of them received data it is consider to be delivered.

7. Conclusion

From past few years' wireless communication is focusing on energy efficiency and reliability, which is desirable in WSNs environment to act as an efficient multicast routing protocol. The automated mobile nodes of WSNs communicate directly with the nodes within its wireless range or indirectly with other nodes in a network. This paper provides a survey of various WSN protocols focusing on energy efficient multicast routing. All the protocol has different mechanism, different assumption to get desired result. So it is very difficult to directly compare these protocols. Review shows that these protocols have different strength and drawbacks. A single protocol cannot fulfill all requirements but each protocol may provide optimized results in a certain given scenario. In designing WSNs, energy efficiency and reliability are the challenging tasks. The proposed solution presented in this paper provides a solution to these challenges. Focusing on energy efficiency and route stability in WSNs, the research will flourish in future and it may be possible to develop a single multicast routing protocol that can meet the entire requirements highlighted above.

References

[1] A Study on Energy Efficient Routing Protocols in MANETs with Effect on Selfish Behavior Aarti Singh, Divya Chadha MMICT&BM, Maharishi Markandeshwar University, Mullana, Ambala, Haryana, India 2012.

[2] S.-J. Lee,W. Su, andM. Gerla, "On-demand multicast routing protocol in multihop wireless mobile networks," Mobile Networks and Applications, 2002.

[3] S. Y. Oh, J.-S. Park, and M. Gerla, "E-ODMRP: enhanced ODMRP with motion adaptive refresh," Journal of Parallel and Distributed Computing, 2008.

[4] J. J. Garcia-Luna-Aceves and E. L. Madruga, "The coreassisted mesh protocol," IEEE Journal on Selected Areas in Communications, 2000.

[5] C.-K. Toh, G. Guichal, and S. Bunchua, "ABAM: on-demand associativity-based multicast routing for ad hoc mobile networks," in Proceedings of the IEEE Vehicular Technology Conference (VTC '00), 2000.

[6] Ivan Stojmenovic, Xu Lin, Power aware localized routing in wireless networks, IEEE Trans. Parallel Distr. Syst. 12 (11) (2001).

[7] Multicasting Techniques in Mobile Ad-Hoc Networks Xiao Chen, Computer Science Department South west Texas State University, San Marcos TX, 2005.

[8] M. Maleki, K. Dantu and M. Pedram, Power-Aware On- Demand Routing Protocols for Mobile Ad Hoc Networks, Proceedings of the International Symposium on Low Power Electronics and Design, , 2002.

[9] D. Pompili and M. Vittucci, "PPMA, a probabilistic predictive multicast algorithm for ad hoc networks," Ad Hoc Networks, 2006.

[10] C.-C. Hu, E. H.-K. Wu, and G.-H. Chen, "OGHAM: ondemand global hosts for mobile ad-hoc multi cast services," Ad Hoc Networks, 2006.

[11] oCast: Optimal Multicast Routing Protocol for Wireless Sensor Networks Lu Su, Bolin Ding, Yong Yang, arek F. Abdelzaher, Guohong Cao and Jennifer C. Hou Department of Computer Science Department of Computer Science & Engineering University of Illinois at Urbana Champaign.2009.

KSII The 12th International Conference on Internet (ICONI) 2020.

A Study on the Response of Cyber Threats Based on ATT&CK to Mitigate Vulnerability in the National Defense Network

GwangHyun Ahn[1], WonHyung Park[2], Ji Won Kang[1], DongKyoo Shin[1]*
[1]Department of Computer Engineering, Sejong University, Seoul, South Korea
[2]Department of Information Security Engineering, Sangmyung University, Cheonan, South Korea
[e-mail: rhkdgus8781@sju.ac.kr, whpark@smu.ac.kr, {jwkang, shindk}@sejong.ac.kr]
*Corresponding author: Dongkyoo Shin

Abstract

Recently, military operations have been affected by, various factors that threaten the defense cyberspace and thru information leakage accidents. Damage to defense cyber assets and the defense information system can result in major disruptions in the execution of current operations. Accordingly, measures are needed to respond to such threats. This paper proposes a plan to efficiently mitigate vulnerabilities in the military internal network by drawing up an attack and defense scenario based on the main weak elements of the military internal network and thru the design of cyber-threat response plan based on ATT&CK.

***Keywords*:** ATT&CK, Cyber Threat, Military Network, Cyber Kill Chain, Vulnerabilites, Mitigate

1. Introduction

Recently, cyberthreats have developed into precision strike-type cyber attack technology with high capabilities. As an example, in 2016, there was a cyber attack case in the domestic/foreign defense sector, where the National Defense Integrated Data Center, the central management center for all military data of the Korean military, had leaks in military data due to internal and external threats[1]. In 2015, there were cases of hacking of military websites, where the personal information of veterans, and military data were leaked by exploiting vulnerabilities in the Pentagon's computer network[2]. In this paper proposes a cyber threat response plan based on ATT&CK by deriving attack and defense scenarios focused on major vulnerable factors in the military network. Effectiveness has the advantage of efficiently mitigating potential vulnerabilities arising from military internal networks.

2. Related Work

2.1 Analysis of ATT&CK and Cyber Kill Chain Models

ATT&CK provides a knowledge base with excellent access to hostile tactics and skills based on real observations. ATT&CK has expanded its Cyber Kill Chain model to derive detailed technologies, including various tactics. In addition, the use frequency of attackers was high, but a defender can also be used. The attacker subdivides the technology based on the tactical goal to achieve and perform the attack based on knowledge-based information derived by generating an attack scenario for each technology. Defenders can create defensive scenarios based on the attacker's tactics,

This work was supported by the National Research Foundation of Korea through the Basic Science Research Program, Ministry of Education, under Grant 2018R1D1A1B07047395.

techniques, and knowledge to establish a defensive strategy for the attacker's position.

2.2 Defense Network Operating Environment (Domestic)

The Korean military's network operation environment is operated within a closed network separated from external networks, centering on the Defense Information and Communication Network and ISP, and information of each military unit is exchanged in non-real-time through the network connection system and data exchange system. Problems include limits in the ability to handle tasks due to the use of separate networks, limits in the collection of real-time information due to the separation of external networks, and the level of cyber security is relatively low due to the use of closed networks[3][4].

2.3 Pentagon Network Operating Environment

The network structure of the U.S. military was designed to connect to external networks with a command control (GCCS) and combat support system (GCSS). In comparison, information of low importance and sensitivity is connected to NIPRNet (Internet Protocol Router Network), which is in operation, and is connected to external networks. This means that both SIPRNet and NIPRNet can be linked to one operating system, and in the case of the U.S. Department of Defense, it is an open structure that operates in conjunction with the Internet[5][6].

2.4 Comparative Analysis of NIST-CSF (Cybersecurity Framework) and K-CSF

NIST-CSF is an improved framework focused on the framework core. The core provides four key elements (function, category, subcategory, information reference) and five functions (identification, protection, detection, response, and recovery) to achieve specific cybersecurity results. K-CSF is a domestic public-private framework established by reference to the existing NIST-CSF. However, only four of the NIST CSF's five functions were derived and established, while the detection function was excluded. This study proposes a cyber-threat response plan using both the detection function that is not derived from K-CSF and the four functions of NIST-CSF[7][8].

3. ATT&CK-based cyber threat response plan design

3.1 Defense Information System Protection Requirements (defense cyber security regulations)

As cyber threats become more intelligent, the Korean military has been organized into five areas of information protection (network, server, terminal, application, and protection) in accordance with Article 24 of the Defense Cyber Security regulations, focusing on protection and control items and requirements. When an information system is built and operated, it receives a security review (technical security, management security, physical security standards) for network protection(19), server protection(36), terminal protection(11), Application Protection(82) ,and protection management(87).

3.2 Working-level Guidelines for Analyzing and Evaluating Vulnerabilities In Civil, Government and Military

When checking the vulnerability of an information security system, the guidelines for analyzing and evaluating the vulnerability of Critical Information and Communication Infrastructure revised by the Ministry of the Interior and Safety were revised and used in the private sector, public sector and defense sector. The civil and public sectors are based on the Information and Communication-Based Protection Act, and the scope of application in the defense sector is based on the defense information systems operated by each military unit defined in the applicable agencies and targets, referring to the Defense Cyber Security regulation and Defense Security Work regulation.

3.3 A Design of Cyber Threat Response Plan for Defense Network Environment based on ATT&CK

In Korea and overseas, cyber security response methodologies based on ATT&CK are mainly derived and improved by private companies. In the field of defense, there is no matter that can be referred to by privately converting the system or framework based on ATT&CK. This study proposed measures to mitigatethe def

ense network vulnerabilities in **Fig 1**.

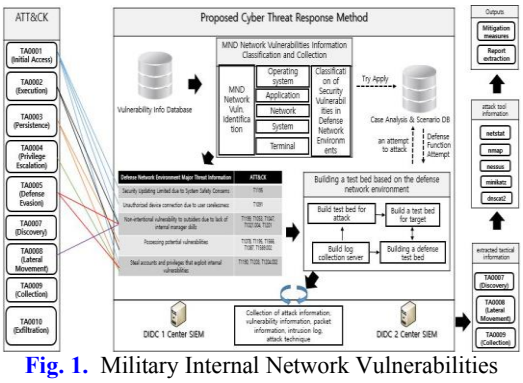


Fig. 1. Military Internal Network Vulnerabilities Mitigation Method

3.4 Research Direction

The cyber threat response method mapped some of the tactics and techniques shown in ATT&CK Matrix with the most common vulnerabilities in actual military units to form and verify an attack and defense scenario. The attack scenario was carried out through attack tools and commands that could exploit vulnerabilities, and the defense scenario produced the SNORT Rule, which resulted in some experiments in which a specific string could be detected, blocked, and mitigated. The goal of this study is to design a framework and a plan that can mitigate vulnerabilities that occur frequently in the defense cyberspace. The cyber threat response measures proposed in this paper are described and show how they can be carried out when applied to the defense network environment. The modeling (draft) design as shown in **Fig. 2** is being carried out, and will be proposed to the extent available in the military.

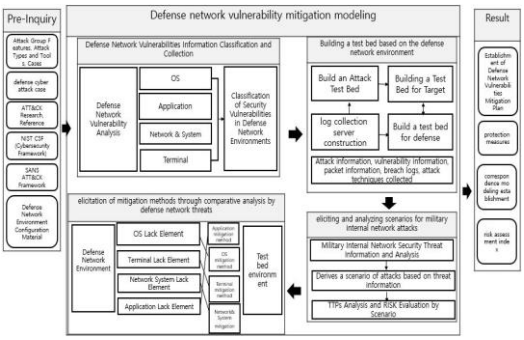


Fig. 2. Defense Network Vulnerability Mitigation Modeling (Proposal)

4. Conclusion

In this paper, we verified countermeasures against cyber threats using defense cybersecurity regulation and ATT&CK data. The study was effective in collecting and identifying information on military internal network vulnerabilities with military specificity in applying ATT&CK information to and research and verify mitigation measures that efficiently protect against cyber-attacks on military internal networks. In the future, we will study cyber security frameworks by establishing a database that can receive information that can be mitigated in real time using ATT&CK Navigator and ELK.

References

[1]Myung-Gil Han, Yong-Hyun Kim, "A Study on System Construction Plan for Cyber Warrior Training", *Korean Journal of Information Protection*, Vol 26, NO.2, pp. 553-540, 2016.

[2]ETH Zurich, "CSS CYBER DEFENSE PROJECT", Vol 1, 2018.

[3]Ministry of National Defense, "Defense Cybersecurity Directive", No 2234, 2016.

[4]Jeon Byung-wook, "A hierarchical approach to improving and protecting the Defense Information and Communications Network", *A Special Paper on Defense Policy Studies at the Korea Institute for Defense Analysis*, pp. 1-26, 2003.

[5]U.S Army, "The Next Generation Theater Geospatial Database", 2006.

[6]Ministry of National Defense, "A Study on the Real-Time Linkage and Integrated Operation of Electronic Government Integrated Network and National Defense Computer Network", 2007.

[7]NIST, "Framework for Improving Critical Infrastructure Cybersecurity", 2018.

[8]Kwon Sung-moon, "Planning the Infrastructure Cybersecurity Framework", *Journal of the Korea Society for Information Protection*, Vol 27, No.2, pp. 241-250, 2017.

[9]KISA, "Detailed Guide to Analysis of Technical Vulnerability of Major Information and Communications Facilities", 2018.

[10] Ministry of National Defense, "Defense Information Protection Directive", No.1109, 2009.

KSII The 12th International Conference on Internet (ICONI) 2020.

Empirical evaluation on a multipath routing protocol for low-duty-cycled WSNs

Hyo Won Cho and Hyungkeun Lee
College of Software Convergence, Kwangwoon University
Seoul, Korea
[e-mail: carl95@naver.com , hklee@kw.ac.kr]
*Corresponding author: Hyungkeun Lee

Abstract

We present the design of a routing metric for an energy-efficient and low-delay routing protocol in wireless sensor networks with a low duty cycle. The new routing metric, EDW, can reduce the energy and delay of transmission paths, which represents total waiting time from the source to the destination due to duty-cycling. In this paper, we propose a new multipath routing protocol based on cross-layer information utilizing the new routing metric, and simulation results show that the proposed protocol shows better performance of the end-to-end delay and the energy consumption as well as the network lifetime.

***Keywords*:** Wireless sensor networks, Cross-layer design, Multi-path routing protocol, MAC, Duty cycle

1. Introduction

Wireless sensor networks (WSNs) consists of a large number of tiny sensor nodes which have limited battery, computational power and storage. These sensor nodes detect an event and transmit data to the sink which works as the gateway to the Internet. Energy-efficiency is one of most significant issues when designing protocols and devices for WSNs [1]. In order to guarantee a long network lifetime, the widely-used approach for saving energy for WSNs is duty cycling. The duty cycling reduces energy consumption by switching sensor nodes between wakeup and sleep states [2].

For duty cycling, sensor nodes in sleep states have their radios turned off and will not be able to transmit data or overhear transmissions, where sensor nodes alternate sleep and wakeup states periodically. The sleep/wakeup scheduling for duty cycling is based on sensor nodes' decision in a distributed fashion. It should guarantee sending and receiving nodes to be active at the same time in order to make data exchange feasible even when seneor nodes operate with a low duty cycle [3, 4].

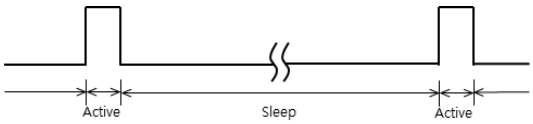


Fig. 1. Duty Cycling

The energy constraint along with the deployment of a large number of sensor nodes causes research challenges in designing MAC and routing protocols for WSNs, where sensor nodes are located far from the sink and therefore use the intermediate nodes to route data toward the sink.

The design goals of a routing protocol for wireless networks are minimizing energy,

This research was supported by the MIST, Korea, under the National Program for Excellence in SW supervised by the IITP (2017-0-00096)

minimizing latency and maximizing throughput. Routing is a selection process of the path between source and destination upon request of data transmission, which is usually accomplished by multi-hop relay. There are two types of routing techniques, single path routing and multiple path routing. Single path routing is simple and scalable, but does not efficiently satisfy the requirements of resource constrained WSNs. It is simple because the route between the source and the destination can be established in a specific period of time. A multipath routing protocol exploits the path diversity provided by multipath to prolong network lifetime by distributing network traffic over multiple disjoint paths [5]. It is a scalable and load-balancing protocol as the networks's size and complexity increase.

When an event occurs in the network, a sensor node in the event area is selected as the source node and the selected source node transmits multiple route-request messages to its neighboring nodes. These route-request messages include different path IDs to construct multiple disjoint paths from the selected source node towards sink. During the route discovery process, all the intermediate nodes select one of their best next-hop neighboring nodes towards sink.

There are two types of duty-cycle scheduling managed by MAC. In the synchronous MAC protocols such as S-MAC [6], the wakeup states of all nodes in the network are synchronized by exchanging periodical control frames. It causes additional control frames leading to waste of energy and redundant collisions between frames. In contrast, asynchronous MAC protocols, such as X-MAC [7], have higher energy efficiency than synchronous protocols. These protocols are using preambles instead of additional synchronized control frames. Because of less control frames and low likelihood of collisions, asynchronous MAC is more effective in terms of energy consumption, while it is quite difficult to match wakeup states between the sender and receiver.

The rest of this paper is organized as follows. Section 2 describes X-MAC and EEMRP based on a new metric, EDW. In the Section 3, experiment results are presented and the performance of the proposed approach is evaluated in terms of delay and energy consumption. Finally, concluding remarks will be given in section 5.

The rest of this paper is organized as follows. Section 2 describes X-MAC and EEMRP based on a new metric, EDW. In the Section 3, experiment results are presented and the performance of the proposed approach is evaluated in terms of delay and energy consumption. Finally, concluding remarks will be given in section 5.

2. Cross-Layer Design of MAC/Routing Protocol

2.1 X-MAC

X-MAC [7] is based on clear channel assessment (CCA) and low power listening (LPL) for accessing channel. Every node with duty cycling has two states: the wakeup and sleep states. Nodes in the wakeup state are active to receive or transmit data, and nodes in the sleep state are inactive to save their power. The key features of X-MAC are the short preamble sampling and the early-ACK mechanism. It transmits multiple short preambles to make receiver to notice. If the receiver detects a short preamble during its wakeup state, it transmits an early-ACK frame to the sender and the sender knows that the receiver is ready. X-MAC saves energy due to the short preamble sampling and the early-ACK mechanism.

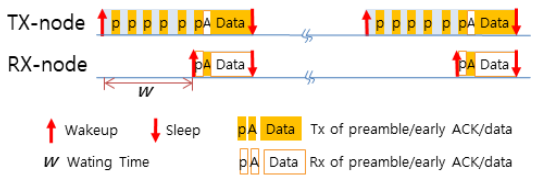


Fig. 2. Operations of X-MAC

In **Fig. 1**, the basic operations of X-MAC are shown, where TX-node keeps transmitting short preambles when TX-node has data to be transmitted to RX-node. Upon receiving an early acknowledgement from RX-node, TX-node starts to transmit data frames immediately. Every node periodically wakes up at its own schedule to check if there are any incoming short preambles intended for it.

While the asynchronous duty-cycling can alleviate protocol overheads, but the waiting time for transmission increases due to different wakeup times of sensor nodes. If this waiting time is accumulated from the source to the destination, a high transmission delay may occur. Moreover, sensor nodes continuously transmit short preambles while waiting for the wakeup at the receiver, which causes a lot of energy consumption.

2.2 Routing metric on waiting time

The most popular metric for routing process is the expected number of transmissions (ETX), with that the path is selected to minimize it required to forward a message to the destination [8]. However, ETX may not be a good metric for routing in asynchronous low-duty-cycled WSNs, since the waiting time for a receiver's wakeup can vary due to the duty-cycle scheduling. Minimizing the waiting time caused by duty cycles can significantly save energy and reduce delay in WSNs.

In order to select the route with minimum waiting time, we introduce a new routing metric for WSNs, expected duty-cycle wait (EDW), that focuses on delay instead of high throughput. It can be achieved by the cross-layer interactions across the physical, MAC and network layers with information of neighboring nodes' s duty cycles. The waiting time due to duty cycles as a routing metric may matter since it reflects the energy consumption and latency of the communication in duty cycled WSNs. Therefore, minimizing EDW of the route can reduce the energy consumption and increases network lifetime.

In duty-cycled WSNs, every sensor node decides and shares its own sleeping schedule, which are usually asynchronous. A sensor node is either in the wakeup state or the sleep state at a given time, and a sensor node has to transmit a message only when its receiver is in the wakeup state. We define the waiting time over a hop, $W_{ij}(t)$, as the time from the moment that the sender i has a packet being ready till the moment that the receiver j goes to the wakeup state which is equivalent to T_{TR} as shown in **Fig. 1**. Therefore, the total waiting time over the path, W_p, is defined as

$$W_p = \sum_{all\ hops} W_{ij}(t)$$

The waiting time in low-duty-cycled WSNs is usually much larger than any other delay, such as processing, transmission, or propagation delay. In this section, we present a new metric, EDW, for routing decision in low-duty-cycle sensor networks. For simplicity, we consider identical asynchronous wakeup schedule, where every node has the same cycle, C.

The routing protocol tries to find the optimal forwarder for each intermediate node on the path so that waiting time due to duty cycles becomes minimum. Let us assume that wakeup time of node N is denoted as t_N in a cycle of C, and data delivery should be postponed to the next cycle when the wakeup time of the receiver is larger than one of the sender. As shown in Fig. 3, S is not able to deliver a message to A within a cycle since the wakeup time of S is less than that of A, while S can deliver a message to B within a cycle. As a result, the path from S to D through B and C has lower latency even though it experiences more transmissions. In the case that cycle length, C, is 16, the total waiting time, W_p, becomes 15 when A relays a message. However, W_p can be reduced to 8 if B and C are selected as intermediate nodes in the example of **Fig. 3**.

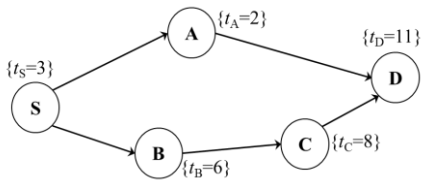


Fig. 3. Example of route selection based on EDW

In traditional non-duty-cycled networks, a sender can send its message to a receiver whenever message is ready. Without considering link failure and transmission collision, the end-to-end delay is proportional to the path's length or number of hops. The most widely-used metric, ETX, has been a metric for a low delay or high throughput path in conventional wireless multi-hop networks.

2.3 Multipath routing protocol based on EDW

Multipath routing protocols for WSNs are responsible for finding disjoint multi-hop paths from source to destination and have to ensure

reliable communication with low energy consumption and latency. The multipath routing has been explored for two reasons. The first is load-balancing: traffic between a source-destination pair is split across multiple (partially or completely) disjoint paths. The second use of multipath routing is to increase the likelihood of reliable data delivery. In these approaches, multiple copies of data are sent along different paths, allowing for resilience to failure of a certain number of paths. Both these uses of multipath are applicable to WSNs.

A multipath routing protocol discovers multiple disjoint paths using a cost function depending on the energy levels and hop distances of the nodes and allocates the traffic rate to each selected path [5]. This is designed to mitigate the energy efficiency requirement of WSNs, while provides reliable data transmission through maintaining a backup path from each source node towards sink.

When an event occurs, a sensor node that detected the event is selected as the source node. Multipath construction is initiated by the sink in WSNs, which is flooding a HELLO message. After initialization, all sensor nodes can obtain information of neighboring sensor nodes and routing information such as the number of hops to the sink and transmission delay. The HELLO message can be broadcast periodically for accuracy.

Sequence No.	Msg. Type (HELLO)	Sink Node ID	Forward Node ID	Hop Count	Cumulative Wait Time

Fig. 4. HELLO message format

In the HELLO message, the sequence number is generated by the sink to distinguish it from other HELLO messages. The message type indicates the HELLO message. The sink node ID and forward node ID indicate the addresses of the sink and the preceding sensor node, respectively. Finally, the hop count and the cumulative wait time record the number of hops from the sink to the current sensor node and the waiting time for duty cycle.

The proposed multipath routing algorithm is based on the protocol [9]. It has the following criteria in order to select the optimal transmission path based on the information of sensor nodes to select an intermediate node.

1) Minimum waiting time (EDW) and

2) Minimum transmissions (ETX)

The sensor node receiving HELLO message transmits an R_REQ (Route Request) message to the sink in order to setup multipath to the sink. The message includes routing information (source and sink nodes information, etc.) and routing metrics (total waiting time and total number of transmissions).

Sequence No.	Msg. Type (R_REQ)	Source Node ID	Sink Node ID	Path ID	Forward Node ID	Path Cost (EDW)	Path Cost (ETX)

Fig. 4. R_REQ message format

The sink receiving an R_REQ message transmits an ASSIGN message to the sensor node through each multipath by limiting the priority of the multipath or the number of paths. Intermediate nodes receiving this message update their routing information and sensor nodes can obtain the expected transmission delay and the number of transmissions to the sink.

The main advantage of this protocol is to increase network lifetime by offloading network traffic over multiple paths according to the cost of data transmission over these paths. The residual battery level of the sensor nodes and their distance to the sink node are considered as the routing metrics. However, a new routing metric can be applied to improve the energy-efficiency and latency in low-duty-cycle WSNs.

Therefore, the information on the wakeup and sleep times of the neighbour nodes can be transferred from MAC to the routing protocol used to alleviate tramsmission delay and energy consumption. If the routing protocol selects a transmission path that minimizes the waiting time due to duty cycling from the source to the destination, it can reduces the overall transmission time as well as energy consumed by sensor nodes.

3. Empirical Performance Evaluation

We run a network simulator in order to evaluate the performance of cross-layer routing protocol with X-MAC. We also show how much a new routing metric has improved in end-to-end latency from sensor nodes to the sink. It is

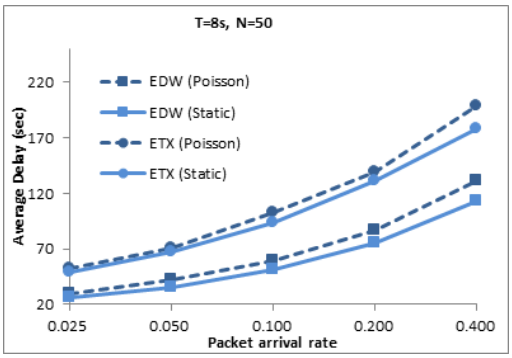
assumed that each node has an omni-directional transmission range where the radio propagation model of two-way ground reflection is employed.

To evaluate the performance of our proposed protocol, we measured energy consumption, and the end-to-end delay under randomly deployed nodes. In WSNs, sensor nodes cooperatively monitor environment to detect events of interest which are highly correlated in space and time. Once an event is detected in the sensed area, multiple nodes around the event area might sense it and try to report data to the sink. Therefore, a synchronized burst of transmissions may be observed which causes severe contentions and collisions at nodes in the event area. Considering the property of sensor networks in the simulation, we exploit random correlated event (RCE) traffic to evaluate the performance of the proposed protocol [10].

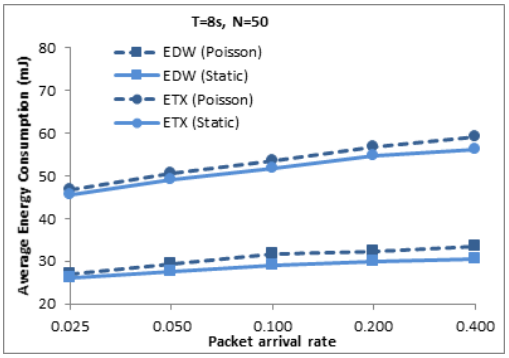
Fig. 4(a) shows the average end-to-end delay of protocols such as a routing protocol based on EDW incorporated with X-MAC and a routing protocol based on ETX with X-MAC, where two types of traffic are applied. One is CBR static traffic, another is Poisson random traffic. From the results, we observe that the routing metric, EDW, improves performance by more than 30% regardless of traffic.

In **Fig. 4**(b), the results of total energy consumption are shown. Since a high level of energy is consumed to transmit short preambles until the receiver wakes up, X-MAC consumes a great portion of energy in sending preambles. In particular, a large amount of energy is wasted due to a number of short preambles in case that waiting time for receiver's wakeup is large in X-MAC. Furthermore, about 55% of energy can be saved by using a metric, EDW, instead of ETX. As a result, we can see that our proposed protocol is effective in terms of energy consumption under different traffic.

Since energy consumption in a node is highly related to the lifetime of the network, **Fig. 5** shows the percentage of residual energy and percentage of alive nodes in the network. From the results, it is shown that the routing metric, EDW, augments the lifetime of the entwork about 20% Moreover, EDW makes sensor nodes to stay alive longer about 15% .



(a) Average delay vs. Arrival rate



(b) Average energy consumption vs. Arrival rate

Fig. 4. Delay/Energy Consumption Comparison

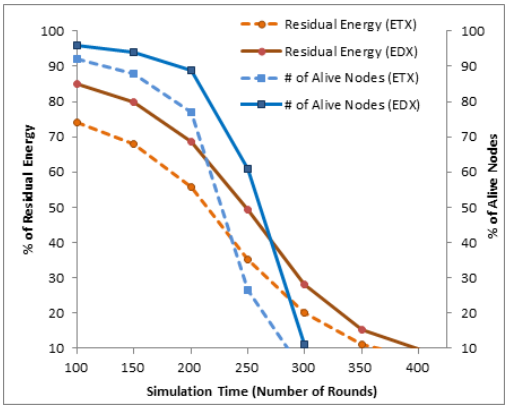


Fig. 5. Residual energy and Number of alive nodes

4. Conclusion

This paper has presented the design and evaluation of a cross-layer routing protocol utilizing a new metric over X-MAC for sensor networks with asynchronous low duty-cycle. Energy saving has been a research issue for WSNs since the lifetime of networks is critical.

X-MAC has been designed to minimize energy consumption by using short preambles. In addition, a routing protocol exploiting sleeping information provided by MAC can reduce latency remarkably.

We conducted simulation to evaluate the performance of the EDW-based routing protocol with X-MAC under various traffic. As a result, the proposed protocol reduces unnecessary preambles, and also decreases latency by lower waiting time due to a new routing metric. Simulation results also show that our proposed protocol outperforms the previous protocols in terms of network lifetime and the number of alive nodes.

References

[1] J. Yick, B. Mukherjee and D. Ghosal, "Wireless sensor network survey," *Computer Networks*, Vol.52, No.12, pp. 2292-2330, August, 2008.

[2] K. Han, J. Luo, Y. Liu and A. V. Vasilakos, "Algorithm Design for Data Communications in Duty-Cycled Wireless Sensor Networks: A Survey," *IEEE Communications Magazine*, Vol.51, No.7, pp. 107-113, July 2013.

[3] Y.Li, W. Ye, J. Heidemann, "Energy and Latency Control, in Low Duty Cycle MAC Protocols," in *Proc. of IEEE Wireless Communication and Networking Conference*, March 2005.

[4] Q. Cao, T. Abdelzaher, T. He, and J. Stankovic, "Towards optimal sleep scheduling in sensor networks for rare-event detection," in *Proc. of International Symposium on Information Processing in Sensor Networks*, pp. 20–27, April 2005.

[5] Y. M. Lu and V. W.S. Wong, "An Energy-Efficient Multipath Routing Protocol for Wireless Sensor Networks," *International Journal of Communication Systems* Vol.20, No.7, pp. 747-766, July, 2007.

[6] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient MAC protocol for wireless sensor networks," in *Proc. of International Annual Joint Conference of the IEEE Computer and Communications Societies*, pp. 1567-1576, June 2002.

[7] M. Buettner, G. V. Yee, E. Anderson and R. Han, "X-MAC: a short preamble mac protocol for duty-cycled wireless sensor networks," in *Proc. of ACM Conference on Embedded Networked Sensor System*, pp. 307-320, Nov 2006.

[8] S. J. Douglas, D. Aguayo, J. Bicket, and R. Morris, "A High-Throughput Path Metric for Multi-Hop Wireless Routing," *Wireless Networks*, Vol.11, No.4, August 2003.

[9] Sunghune Lee, Hyungkeun Lee, "Research of Energy-efficient Multipath Routing Algorithm in Wireless Sensor Networks," *Journal of the Korea Society of Computer and Information*, Vol. 14, No. 8, pp. 59-64, 2009.

[10] R. Shakya, Y. Singh, N. Verma, "A correlation model for MAC protocols in event-driven wireless sensor networks," in *Proc. of IEEE TENCON*, pp. 1–6, Nov. 2012.

Clustering Techniques to Reduce Routing Overhead in Ad hoc Networks

Hawezi Omeed kamal khorsheed, MD Asri Bin Ngadi, Abdul Hanan Abdullah and Israr Ghani
1Faculty of Computer Science, University Technology Malaysia (UTM), Malaysia
Skudai, Johor Bahru, 81310, Malaysia
Omeed.kamal-1984@graduate.utm.my, hanan@utm.my, dr.asri@utm.my, israrghani@outlook.com
Corresponding author : omeed kamal

Abstract

Information and communication technologies have changed the data communication operations and systems. Ad hoc data communication has gained attention due to its flexible infrastructure less or infrastructure based network requirements. Routing protocols have significant role to establish the communication in these networks. Due to flexible features, ad hoc networks have suffered with various routing, limited resources, unpredictable and dynamic topologies, overhead and mobility challenges. In order to address these challenges' different types of approaches have been designed and clustering is one of them. In this paper, we present clustering approaches for ad hoc networks, its requirements, process and formation. In addition, we also discuss clustering protocols designed for four domains. The systematic comparison also presents in this paper. This paper will help to guide new researchers for further investigation in ad hoc networks domains.

Keywords: Ad-Hoc, Wireless Network, Clustering, Distributed, Multi-Hop, Mobility, Routing Table

1. Introduction

Ad hoc networks refer to special and unique type of data communication with or without any preexisting infrastructure. Most of the advance fields have successfully adopted ad hoc networks for data communication such as mobile computing, intelligent transportation, agriculture, industries and healthcare and monitoring systems. In ad hoc networks, different types of nodes are establishing a temporary wireless networks without any administration or prior configuration [1]. The nodes are formed using dynamic link and not rely on existing infrastructure such as gateway, and router. Every node is capable to sense, collect, store and further forward the data using peer-to-peer networking. Nodes in these networks are static or dynamic with mobility features to change their location and configure itself to communicate for various purposes. These networks provide cost effective solutions, short range and long range communication capabilities among different heterogeneous network units. User can collect the sense data which has valuable and trust based information. Inter of Things (IoT), Wireless Sensor Networks (WSNs), Vehicular ad hoc Networks (VANETs), Underwater Sensor Networks (UWSN), Body area Networks (BAN) are examples of heterogeneous, multi-hop networks. With various beneficial features, ad hoc networks have suffered with different challenges such as dynamic topologies, mobility of nodes, interference, connectivity, data security, noise, overhead and delay [2]. The routing protocols are essential to address various data communication routing. Different architectures and protocols have designed for self-organizing and construct temporary ad hoc networks with strong robustness.

Routing protocols have designed to select the optimal path in a network in order to send data over the network. To find the optimal path is another challenge in ad hoc networks due to infrastructure less environment, unstable

topologies and other limited resources. Sometime, suddenly any node can disappear or become out of the service and lead to disconnectivity and delay issues in networks. In order to address these limitations, different types of location based, topology based and clustering approaches have been developed [3]. Data transmission depend on the nodes transmission range of source node and its neighbor nodes. If the sender or source node finds the destination node within its radio transmission range, then directly send the data to destination node. On the other hand, source node selects the next forwarder nodes node as a router pass the data packet to the destination node. Clustering is one of the well-known routing technique to arrange the network member nodes into a hierarchy of small groups depending on a specific consideration. Clustering is providing network scalability, routing faller and network management in ad hoc networks.

In this paper, we discuss the different routing issues of ad hoc networks and their routing solutions. This paper presents a depth overview to new researchers of clustering approaches in ad hoc networks adopted for different domains.

The rest of the paper is organized as follows: Section 2 discusses the ad hoc network functions, communication standards and architecture. Section 3 presents the routing protocols and their limitations. Section 4 illustrates the clustering protocols, their cluster head selection process, formation and management. Section 5 presents the discussed clustering protocols comparison and in-depth overview. Last section concludes the paper with future direction.

2. Ad hoc Networks

Ad hoc networks are mostly based on regular sensor nodes and other communication devices for data communication. These networks have equal opportunities for monitoring and sensing the data and forward for further processing and analysis. The nodes distribution is based on their resources where the regular nodes are responsible for monitoring the enlivenment and special nodes including gateway, access points and router are responsible for collecting the data and pass to data centers [4]. Due to environment impact and limited energy, the sensor nodes are prone to failure and create complexities in data routing process. On the other hand, the VANETs has suffered with high mobility of nodes and location identification issues. Another issue is changing topologies with moving vehicle nodes and has serious accuracies impact in network. Ad hoc networks are self-organized networks in which the distributed and center free network architecture maintaining the connections. Sensor nodes are randomly join and leave the network and cause of unpredictable network topologies. Another issue in these networks is limited bandwidth especially for high data rate applications.

Energy is another significant constraint of ad hoc networks especially in WSN, WBAN, UWSN. Due to heterogeneous capabilities, the sensor nodes are depleted early and disturb the normal data transmission flow in the network [5]. Energy is more important criteria especially where the replacement and charging of nodes is not possible such as volcano monitoring, rainforest and desert applications. **Fig. 1** shows the ad hoc network types.

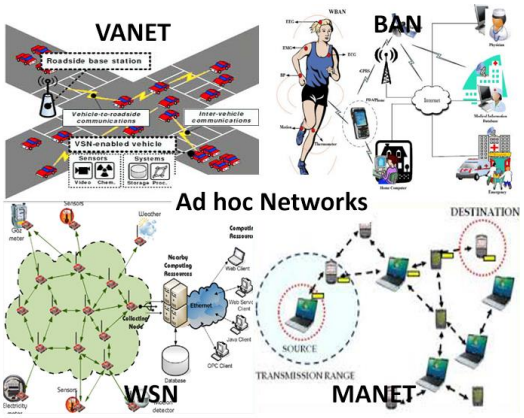


Fig. 1. Ad hoc Networks

3. Routing Protocol for Ad hoc Networks and Requirements

Different types of routing protocols have been designed for reliable data communication in ad hoc networks. The routing protocols are required to handle limited resources of ad hoc networks such as traffic density, dynamic topologies, network partitioning and interference. Routing is method to find the data path between source and destination node. The design of routing protocols for ad hoc networks are affected by various factors. Routing protocols have important role in

energy depletion because overhead and delay issues consume more energy in the network. With the increase of ad hoc networks on large scale, the networks are facing serious routing challenges. The most feasible routing protocols are needed to handle the self-organizing networks to ensure the topological changes. The current routing protocols are difficult to handle and fulfill the ad hoc networks requirements. Ad hoc networks have been embedded with various types of networks such as mobile networks, sensor networks and wireless networks. Thus, the new protocols should have strong routing process to handle advance ad hoc networks. There are some important factors to design the routing protocols for ad hoc networks such as node deployment, energy consumption, node nature, node coverage, quality of services requirements and scalability. Node deployment is always depending on application type and used routing protocols. The node deployment is random or randomized where node have deployed and routed by predefined paths [6].

Energy utilization or consumption is another main factor of routing protocol design in efficient manner. Sensor nodes consume energy during processing, monitoring, receiving and transmitting the data. Since, the sensor nodes have limited capabilities and complex processes consume more energy and cause of node depletion, connectivity changes, reorganization and packet loss. The routing protocols must have strong tradeoff between sensor operations, accuracy and energy optimization. Coverage area is also very essential for design a routing protocol because each sensor node prevails a certain view of environment. There are different parameters to fulfil the quality of services requirements in ad hoc networks such as bandwidth, data delivery, jitter, throughput and delay. Most of the ad hoc networks applications are based on real based data monitoring and time sensitive data. The routing protocols are application specific such as data driven, query driven and event driven methods. In time driven methods, sensor node reports the collected data whenever event occur in the network.

4. Cluster based Routing

There are various different types of routing protocols have been proposed for ad hoc networks such as topology based, position based, reactive, proactive, next hop selection, single hop, multi hop and cluster based. Through literature, we found that clustering approach has many advantages over state of the art routing types [7, 8]. Clustering approaches have addressed energy consumption issues, pure routing issues and network formation issues by using energy efficient methods. These protocol always selecting the cluster head nodes having high energy resources for sensing and processing. This feature provides scalability, energy minimization and lifetime maximization. In this section, we discuss latest protocols and their features for different domains.

Clustering protocols enhance the network scalability and obviate routing faller with more efficient management system and reduces the transmission overhead. Clustering divides the sensor nodes into four main positions: ordinary nodes, secondary nodes, cluster head and gateway nodes. The ordinary nodes are normal nodes present in cluster where Cluster Head (CH) node is responsible to communication at Inter-Cluster level (members neighbor table inside the same cluster group) and Intra Cluster level (all cluster head nodes with other cluster heads) [9]. The secondary nodes help to the main CH when it fails to keep a backup of routing and cluster information. All Ordinary nodes can serve as secondary CH, and only CH node can authorize ordinary node as a secondary cluster head node. Gateway nodes or undecided is very vital routing node, it stores the gateway table, when CH node sends the RREQ packet (message) to its own gateway members nodes. The gateway members node will forward the data packet to all CHs which are existing in its gateway table. Gateway nodes in other cluster groups will resend the RREQ packet until it reaches the destination node or any intermediate node within a route to the destination node. Each cluster may have more than one gateway node. **Fig. 2** shows the cluster nodes types and positions.

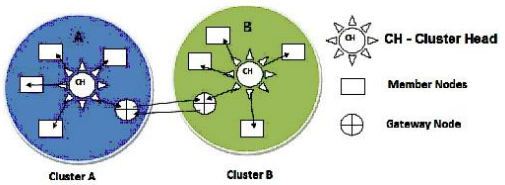


Fig. 2. Clustering Nodes

4.1 Mobile ad hoc Networks

There are various clustering protocols have been designed for MANET. Energy is always a challenge in MANET and most of the protocol have addressed the energy issue by different methods. Rajesh, et al. [10] proposed a novel cluster depended energy efficient routing protocols based on clustering method. In this protocol, the CH has selected randomly time-to-time to enhance the network lifetime. The CH selection is based on the comparison of effective available energy to increase the life time of network. Amagata, et al. [11] proposed an efficient query routing protocol for MANET. In this protocol, a new clustering method has adopted based on query processing where the nodes holding high rank data for CH selection. Each CH maintains a set of hop counts between itself and other high rank data items. It further judges about data transmission. In addition, this protocol achieves top-k query processing by only necessary nodes. This protocol easily adopts topological changings and ensure accuracy for highly dynamic networks. Clustering methods also have adopted to enhances the trust and security in MANET. **Fig. 3** shows the protocol process.

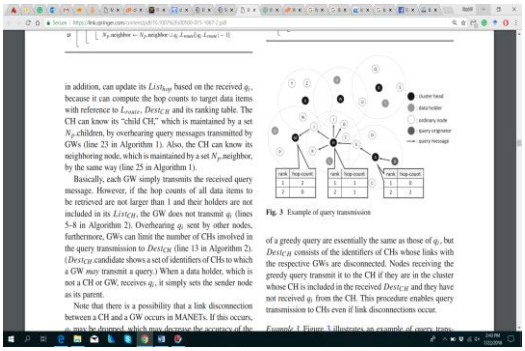


Fig. 3. Cluster formation, cluster heads and routing process

Another clustering protocol based on ant optimization clustering method was designed in [12], for MANET. This clustering approach is focusing the trust values and addressed jellyfish attacks in MANET. Jellyfish attack refers to widely well-known vulnerability of TCP where the route is changing by means of applying multipath routing methods. This protocol depends on ant forging behavior and shortest path has identified by traversing the simple route. The request packets broadcasted by ordinary nodes to CH then it forwards the request to destination node or adjacent CHs. There are many cluster based protocols for secure the data in MANET such as [13-15].

4.2 Wireless Sensor Networks

With new advancements, WSN is one of the prominent filed of ad hoc networks based on spatially distributed independent sensor nodes. Cluster based routing also has implemented in WSN and has various positive features to reduce the energy, and routing issues in the network. An energy cross layer adaptive threshold routing protocol was proposed in [16] using adaptive threshold distributed method. The network model has used for formation the CH selection and radio energy model focuses on calculating the energy during data communication. The proposed network model has a square field based on uniformly distributed and random sensor nodes. In addition, every sensor node become steady and accumulate the data from field and bypass the base station. The communication between ordinary sensor nodes and CHs is possible using single-hop communication. The relative distance is approximated using input of signal strength data. Yan and Wang [17] proposed an adaptive clustering protocol using neighbor node energy level for WSN. This protocol dynamically adjusts the CH selection by using residual energy and average energy level. This protocol scales up the CH selection and also considering the neighbor nodes density and also difference between residual energy of nodes.
Sreevidya and Rajesh [18] proposed an enhanced energy optimized cluster based protocol for WSN. In this protocol, the CH is selected for route formation and minimizes the energy consumption in WSN. The CH selection with more residual energy and then rotate the CHs periodically based on probability. The rotation policy will ensure the distribution of energy for network lifetime. After CH selection, a broadcast message is passed to all other ordinary nodes within the cluster. Then CH identifying the path to sink node using other CHs. This protocol is on demand in nature. The rotation is used to enhance the network lifetime and probability of node which has high chance as a CH. Awan, et al. [19] proposed an energy efficient cluster

routing for WSN using unequal sized clusters of sensor nodes to transmit the data to base station.

4. 3. Body Area Networks

Special kind of sensor nodes form a network by putting the biomedical sensor nodes in or on the patient body. Due to limited resources, the nodes are not able to deliver the serious data towards the medical centers and hospitals. The WSN routing protocols are not directly implemented in WBAN due to scalability issues. Ali and Khan [20] proposed an energy efficient cluster based protocol for BAN. This is hybrid security framework by using multiple clusters. This protocol also addresses the energy issue by using the electrocardiogram based key method. Energy consumption is calculated through distance between ordinary and CH nodes and security by using PV-based pairwise key generation. The proposed protocol provides secure cluster formation where pairwise management has adopted for each data communication among sensor nodes. Sharma, et al. [21] proposed a clustering to minimize the energy consumption for WBAN. In this protocol, the sink is placed in center of human body and multi-hop data transmission has utilized to elect the CH. The CH selection method is threshold technique where the previous CH used for next round. Author claimed that by using this method the energy is not wasted due to repeating the CH election. If the CH has less energy calculated by threshold value, then it replaced by new CH with more energy resources.

Cluster based Energy Efficient Routing Protocol proposed in [22] by using base station as a gateway in WBAN. This protocol focuses on sink node placement at center and sensor nodes distributed in homogenous mode to determine the energy level at initial stage. The energy level is based on transmission distance and energy consumption in the network. This protocol forms a cluster and then elect a CH and static base station outside the network to address the energy issues.

4.4. Vehicular Ad hoc Networks

One of the unique characteristics of VANET is high mobility which leads to excessive overhead and link disconnections. Clustering methods have been designed to address the mobility and frequent changings issues in VANET. In this context, control overhead reduction cluster based routing protocol was designed by [23]. In this protocol, the network is divided into multiple clusters and addressed the overhead issue by calculating the optimal period for updating the control messages between ordinary cluster members and CHs. Author evaluated the proposed protocol with state of the art protocols and indicated that protocol reduces the continuously broadcasting the hello messages and generate minimum packets. Qureshi, et al. [24] proposed a Cluster-based Routing for Sparse and Dense networks to handle high mobility and dynamic topologies. The cluster formation in this protocol is based on vehicle node distribution in the network. Cluster formation stables and balance the routing by using duty cycle and periodic re-clustering methods. After formation, this protocol elects the CH with less overhead node.

Suma and Lalitha [25] proposed a hybrid protocol for data dissemination in VANET. This protocol combines the hierarchical clustering method and priority based messaging. The CH communicates with other ordinary nodes and only active nodes are allowed for data communication. To check the probability of transmission, multinomial probability distribution is used and providing the probable messages transmission.

5. Comparison

Table 1 summarizes the comparison of cluster based routing protocols in four domains of ad hoc networks on the basis of energy, security, complexity, overhead and delay. The selected protocols are latest clustering protocols and from four different domains. The data communication in these protocols is based on pure ad hoc or with infrastructure. MANET, WSN and BAN have energy issues due to limited resources of sensor nodes where the charging and changing of batteries are not possible or difficult [26]. On the other hand, the VANET protocols address the pure routing issues such as overhead, delay and disconnectivity issues. Complex routing strategies are creating more overhead in the network and cause of data delay. In discussed routing protocols, some protocols have less overhead and some have more overhead due to their complex routing processes to select the best

Table 1. Protocols Comparison

Mobile Ad hoc Networks					
Author/Year	**Energy**	**Security**	**Complexity**	**Overhead**	**Delay**
Rajesh, et al. [10]	✓	×	✓	×	×
Amagata, et al. [11]	✓	×	✓	✓	×
Satheeshkumar and Sengottaiyan [12]	✓	×	✓	✓	✓
Wireless Sensor Networks					
Singh and Verma [16]	✓	×	✓	×	×
Yan and Wang [17]	✓	×	✓	✓	✓
Sreevidya and Rajesh [18]	✓	×	✓	✓	×
Awan, et al. [19]	✓	×	×	✓	×
Body Area Networks					
Ali and Khan [20]	✓	✓	✓	×	×
Sharma, et al. [21]	✓	×	×	✓	×
Alghamdi [22]	✓	×	✓	✓	✓
Vehicular Ad hoc Networks					
Abuashour and Kadoch [23]	×	×	✓	✓	✓
Qureshi, et al. [24]	×	×	✓	✓	✓
Suma and Lalitha [25]	×	×	✓	✓	×

path or next forwarder node in the network [27]. The much research has been done on clustering methods to minimizing the energy consumption and enhance the routing. However, with many advantages' these protocols still have issues such as calculation and selection of CH, scalability, topological adoption, fault tolerance and redundancy management [6].

5. Conclusions

Ad hoc networks have gained attention all over the world due to its broad scope and cost effective solutions for data communication. Routing protocols are using to make the communication possible among mobile and static nodes. Various routing solutions have been presented and through literature we found clustering protocols have more features to handle ad hoc networks issues. This paper discussed selected cluster based routing protocols, their formation, cluster head selection and routing processes. In addition, this paper has covered four main domains of ad hoc networks including MANET, WSN, BAN and VANET and discussed their latest cluster based routing solutions. The in-depth comparison also presented in this paper. This paper will help the new researchers who are willing to work in ad hoc networks domain. In future, we will focus on clustering approaches in IoT and cloud computing.

References

[1] M. B. Yassein and M. Shatnawi, "A Survey of Probabilistic Broadcast Schemes in Mobile Ad Hoc Networks," in Critical Research on Scalability and Security Issues in Virtual Cloud Environments, ed: IGI Global, 2018, pp. 269-282.

[2] K. N. Qureshi, A. H. Abdullah, and R. Yusof, "Position-based routing protocols of vehicular Ad hoc networks & applicability in typical road situation," Life Science Journal, vol. 10, pp. 905-913, 2013.

[3] J. Yu, Y. Qi, G. Wang, and X. Gu, "A cluster-based routing protocol for wireless sensor networks with nonuniform node distribution," AEU-International Journal of Electronics and Communications, vol. 66, pp. 54-61, 2012.

[4] S. K. Sarkar, T. G. Basavaraju, and C. Puttamadappa, Ad hoc mobile wireless networks: principles, protocols, and applications: CRC Press, 2016.

[5] H. Samuel, W. Zhuang, and B. Preiss, "Improving the dominating-set routing over

delay-tolerant mobile ad-hoc networks via estimating node intermeeting times," EURASIP Journal on Wireless Communications and Networking, vol. 2011, p. 7, 2011.

[6] S. P. Singh and S. Sharma, "A survey on cluster based routing protocols in wireless sensor networks," Procedia computer science, vol. 45, pp. 687-695, 2015.

[7] C. Gherbi, Z. Aliouat, and M. Benmohammed, "A survey on clustering routing protocols in wireless sensor networks," Sensor Review, vol. 37, pp. 12-25, 2017.

[8] C. Cooper, D. Franklin, M. Ros, F. Safaei, and M. Abolhasan, "A comparative survey of VANET clustering techniques," IEEE Communications Surveys & Tutorials, vol. 19, pp. 657-681, 2017.

[9] J. Prakash, D. Kumar Gupta, R. Kumar, J. Prakash, R. Kumar, J. Saini, et al., "A survey of clustering schemes for mobile ad hoc networks," Journal of Artificial Intelligence, vol. 10, pp. 264-281, 2017.

[10] M. Rajesh, T. Gireendranath, and J. Murthy, "A Novel Energy Efficient Cluster Based Routing Protocol for Highly Dense MANET Architecture," International Journal of Computational Intelligence Research, vol. 13, pp. 719-744, 2017.

[11] D. Amagata, T. Hara, Y. Sasaki, and S. Nishio, "Efficient cluster-based top-k query routing with data replication in MANETs," Soft Computing, vol. 21, pp. 4161-4178, 2017.

[12] S. Satheeshkumar and N. Sengottaiyan, "Defending against jellyfish attacks using cluster based routing protocol for secured data transmission in MANET," Cluster Computing, pp. 1-12, 2017.

[13] S. S. Jamaesha and S. Bhavani, "A secure and efficient cluster based location aware routing protocol in MANET," Cluster Computing, pp. 1-8, 2018.

[14] A. K. Gupta and S. Prakash, "Secure communication in cluster-based ad hoc networks: a review," in Next-Generation Networks, ed: Springer, 2018, pp. 537-545.

[15] N. J. K. Patel and K. Tripathi, "Trust Value based Algorithm to Identify and Defense Gray-Hole and Black-Hole attack present in MANET using Clustering Method," 2018.

[16] R. Singh and A. K. Verma, "Energy efficient cross layer based adaptive threshold routing protocol for WSN," AEU-International Journal of Electronics and Communications, vol. 72, pp. 166-173, 2017.

[17] A. Yan and B. Wang, "An adaptive WSN clustering scheme based on neighborhood energy level," in Information Technology and Mechatronics Engineering Conference (ITOEC), 2017 IEEE 3rd, 2017, pp. 1170-1173.

[18] B. Sreevidya and M. Rajesh, "Enhanced energy optimized cluster based on demand routing protocol for wireless sensor networks," in Advances in Computing, Communications and Informatics (ICACCI), 2017 International Conference on, 2017, pp. 2016-2019.

[19] K. M. Awan, A. Ali, F. Aadil, and K. N. Qureshi, "Energy efficient cluster based routing algorithm for wireless sensors networks," in Advancements in Computational Sciences (ICACS), 2018 International Conference on, 2018, pp. 1-7.

[20] A. Ali and F. A. Khan, "Energy-efficient cluster-based security mechanism for intra-WBAN and inter-WBAN communications for healthcare applications," EURASIP Journal on Wireless Communications and Networking, vol. 2013, p. 216, 2013.

[21] R. Sharma, H. S. Ryait, and A. K. Gupta, "Clustering Based Routing Protocol to increase the stability in WBAN," Reasearch Gate, 2015.

[22] T. A. Alghamdi, "Cluster based energy efficient routing protocol for wireless body area networks," Trends in Applied Sciences Research, vol. 11, p. 12, 2016.

[23] A. Abuashour and M. Kadoch, "Control Overhead Reduction in Cluster-Based VANET Routing Protocol," in Ad Hoc Networks, ed: Springer, 2018, pp. 106-115.

[24] K. N. Qureshi, A. H. Abdullah, F. Bashir, S. Iqbal, and K. M. Awan, "Cluster- based data dissemination, cluster head formation under sparse, and dense traffic conditions for vehicular ad hoc networks," International Journal of Communication Systems, vol. 31, p. e3533, 2018.

[25] G. J. Suma and R. Lalitha, "Vehicular Ad hoc Networks: A hybrid approach to data dissemination in exigency situations,"

Wireless Networks, vol. 22, pp. 1725-1737, 2016.
[26] S. Singh and S. Agrawal, "VANET routing protocols: Issues and challenges," in Engineering and Computational Sciences (RAECS), 2014 Recent Advances in, 2014, pp. 1-5.
[27] O. Ahmad, "Survey of routing protocols in wireless sensor networks," 2014.

Comparative Analysis of Access Control Models: Types and Challenges

Aitizaz Ali[1], Muhammad Fermi Pasha[2], Ong Hue[3], and Muhammad Hasnain[4]
[1,2,3,4] School of Information Technology, Monash University Malaysia.
[aitizaz.ali@monash.edu, muhammad.FermiPasha@monash.edu,ong.hueyfang@monash.edu, muhammad.malik1@monash.edu]
*Corresponding author: Aitizaz Ali

Abstract

Access control is aimed to control the access of authorized and unauthorized users through system policies. Access control follows predefined rules to allow or reject user access to specific data. These rules differ according to the organization, role, or attributes. This paper presents an analysis and comparison of access control models according to their functions, merits, and demerits. A simple review method is used to review the chosen papers based on access control models. This paper's results showed that researchers often underestimated physical security compared to malware, cyber espionage, and hacking. This paper also identified two types of access control, such as physical security and information security. The challenges related to the data are the use of a proper access control system. The access control criteria are defined in this research with limitations. Identified challenges include access control, implementation cost, flexibility, and fine-grained access control system. Threats and breaches violate physical security, including brute-force attacks. Many common access control models exist based on objects, identity, role, attributes, time, domain, and trust.

Keywords: Access control, polices, security, trust, privacy, cryptography, encryption.

1. Introduction

In the traditional symmetric fundamental model, symmetric key is used to encrypt the data. Data owner data into some groups and then encrypts these groups using the symmetric key. Users who have the key can perform decryption of the data. In this scheme, authorized users are listed in the Access Control List (ACL). The major drawback of this scheme is the linear increase in number of keys with the increased number of data groups [1]. Also, if any change occurs in the user and data owner relationship, it will lead to affect other users in the ACL. So, in summary, this scheme is not in practical use in different scenarios [2].

Standard ACL defines a set of rules on how to prevent from different attacks. In this technique, network traffic is filtered out. The firewall mostly configured using standard ACL. This scheme follows the source IP only. Extended ACL can be called the gatekeeper of a network, Its always implemented at the network edge. The extended ACL decides to grant or deny access based on port number, IP, source, destination, and current time [3].

2. Types of Access Control Models

In this section, we have discussed various access control models in a hierarchical order.

2.1 Discretionary Access Control (DAC) and MAC (Mandatory Access Control)

DAC and MAC are considered as object-based models. The DAC object is directly connected to the subject using the relationship between these two entities. MAC is the improved form of DAC that uses the security attributes of the subject and objects to grant access [6]. MAC (Mandatory Access Control) is considered a standard and well-established approach in cryptography [8]. It was first designed for military purposes for controlling information. MAC is indeed based on a lattice-based in-formation flow model. MAC has further two versions, such as Bell-LaPdula and Biba models. The Bell-LaPadula model provides information flow and confidentiality, whereas the Biba model is concerned with maintaining the data's integrity.

In MAC, there is no concept of ownership. In other words, it can be described that in MAC, user rights and privileges are not resource-centric [9].In order to understand more, the working of the MAC model, partial orders, and lattices must be understood. Partial orders use mathematical sets and set properties. Partial orders modeling in MAC is used to match and order resources and user's attributes properly. A lattice consists of the greatest lower bound and least upper bound set. Lattice is used in MAC when information flow is almost critical. That's the main reason it was purposely designed for the defense sector. MAC uses lattice to follow the information flow policy [7]. Information flow policy deals with the transition of information flow between various security levels. Information flow can be monitored and maintained by assigning each object a security level or class.

The government and military departments mainly used it. Accesses to the resources are controlled by the OS as configured by the system admin. MAC controls labels and security tags on each resource or object. This security information can be classified as secrets, top secrets, and confidential. These security levels show management and access levels to specific information. As a client requires to access a particular object, the OS first checks the user classification according to the object security level. If the user satisfies the object sensitivity level, then the access is granted otherwise rejected.

MAC model provides most of the security, but the overhead is considered to other models. Another drawback of using MAC is its very expensive due to its system over-head. Another scheme that was proposed in the literature is the integration of a public key and asymmetric key. This approach is different from the various other approaches in a light way; it uses a public key to encrypt data.

2.2. Role-Based Access Control (RBAC)

RBAC has been applied using cloud computing. RBAC provides flexibility to access control management at a close level to the organization's policy and format. Secure RBAC was proposed by Zhou et al. using RBE (Role Based Encryption). In RBE, user revocation is supposed to be flexible. This approach provides users who have authorized roles can only decrypt data. Dynamic role-based model approach was proposed with the objective of integrating integrated with the RBAC using cloud computing. Many RBAC approaches have been introduced earlier, including those that mostly can't provide flexible data access demands that depend upon trust.

In another reason is that fine-grained access control can't be supported inside the role [10]. RBAC is easy to manage in the context of role assignment to user dynamically change with time but stable in permission assignments. The main function of the role is it encapsulates lower-level permissions and privileges from a higher level. In other words, it can be concluded that roles are called pro-data-abstraction. The context of roles depends upon the environment, organization, and time. In early studies, the role can be defined as a generalized hierarchy and agent representing users who are assigned a role. Ting defined the application of role for application-level security. Ferraiolo and Kuhn defined the modern definition of RBAC. In the following, we present the types of RBAC.

RBAC is categorised into a) Basic RBAC, b) hierarchical RBAC, and c) constrained RBAC. Basic RBAC is based on a set of users, resources, and a set of the access permissions. The core behind this to integrate subsets of

access rights and permissions within named roles. When a user is assigned a role, it vets the authority in accessing the particular source within the confinement of the role. A role defines competency in a specific area. A virtual role or position in RBAC is used when a role is assigned without any direct user. For example, the role of a health-care provider is considered a high-level abstraction for doctors and nurses. Role assignment represents a concrete scope of responsibility. At a lower level, access to resources is provided through a functional interface by a resource manager. Such an interface to a resource is called operation. It depends upon the semantics of the operation that one or more permission might be added to a role but disjoint. For example, a read and append operation be assigned to a role but disjoint. These two operations can be implied through write operation. Depends upon the access policy, the list of the permission to a particular operation can be further evaluated by the access control system.

RBAC model is used to overcome the complexity and overhead in a worst-case scenario. Assignment of user rights and permission can be best implemented inline using the least privilege principle. Role-Based Access Control has been applied using cloud computing.

2.3. ABE and ABAC

Sahai and Waters [5] proposed the model namely 'Attribute-Based Encryption' (ABE). It was the first new access control method proposed based on attributes. The proposed model details the use of the private key its association with the cipher-text and a set of attributes. Only users who have the match of key associated with cipher-text can decrypt the encrypted message. The Sahai-Waters' primary drawback [12] was the limitation to the general system and the threshold semantics, which is not more expressive. The ABE model provides better mechanism than another one to control and access data in cloud computing.

By using ABE, it is easy for someone to share data according to a well-defined encryption policy without having any prior information of the data recipient. Different approaches are applied using access control, but they are based on a centralized mechanism that is more prone to vulnerability and risk.

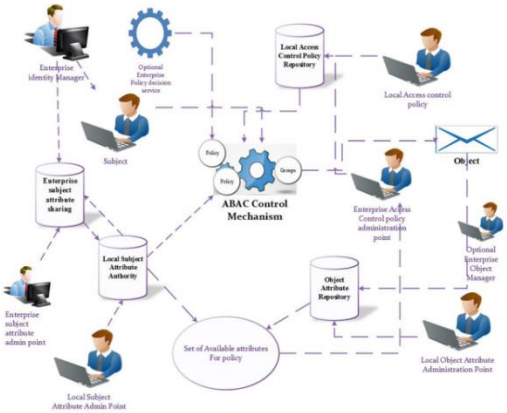


Fig. 1. Working of access control system

Fig. 1 is the illustration of simple working of the 'access control system' (ACS). In which users request for a specific object (data). The request is shown that it needs to be passed through different policies and vendors. After this request needs the condition, then the access is granted. The access depends upon the access control model i.e., ABAC or RBAC or MAC.

The main problem with conventional encryption is that all data are visible to the cloud. The local domain concept arises for patients and users to register to send and receive digital health services. Authentication, and authorization activities along with defined attributes are executed in varying or same infrastructures to maintain the appropriate security levels. The access control system as shown in **Fig. 1** consists of a set of policies and objects that define access control.

2.4. Types of ABE

Goyal et al. proposed a more detailed and general KP-ABE-based encryption system. The construction of Goyal et al. is the extension to Sahai and Waters [7]. In this approach, a cipher-text with users' secret keys and attributes is associated with a set of monotonic structure. Perretti et al. proposed threshold ABE. In this scheme, user revocation was introduced to how it affects the access control. Chase [11] introduced the concept of multi-authority attributes. In this

construction, each authority manages different domains of attributes. Collusion attack is a big issue while designing and implementing multi-authority attributed based access control [4]. More generally, we can define ABE as an Integration of security technique with the ABAC model). Another approach is ABE to accomplish flexibility and fine-grained access control. In the literature, two methods have been proposed using ABE, such as 'Key Policy-Based ABE' (KP-ABE) and 'Cipher-text policy-based ABE' (CP-ABE) [3]. The only difference between these two approaches is how the policies are associated with the cipher-text and decryption keys.

2.5. Hierarchical attribute-based encryption (HBAE)

Hierarchical ABE (HABE) was proposed by [5]. HBAE is the combination of cipher-text based ABE and hierarchical encryption approach. This approach provides features such as flexibility and scalability. Yu et al. [9] proposed a new dimension of ABE access control, which combines two approaches, including KP-ABE and Proxy-Re-Encryption (PRE) [13].In this approach, the data owner is allowed to transfer computational overhead at the server-side. This approach provides user confidentiality and secret key accountability. It takes advantage of fine-grained access control from KP-ABE. Data is encrypted using a Data Encryption Key (DEK). After using DEK for encryption, it is again encrypted using a symmetric key. User revocation [11] is another big issue in the access control approach design. A problem arises when a user is revoked from the system but still have the key. This revoked user has the decryption key so he/she can use it for decryption, which is an issue in most of the proposed approaches. A good strategy is after user revocation, a re-encryption and re-keying process is needed. Data owner can perform re-encryption and re-keying to protect data from the revoked users [10]. A trust-based access model was proposed using trust as metrics for access control [5].

2.6 Trust-Based Access Control Model (TBAC)

Situational TBAC was introduced to provide access according to a location using a trust level [3]. This approach limits the transmission and access of documents according to a location. Performance aware TBAC was proposed to protect sensitive information according to seniority and behavior as a trust level [11, 13]. In order to grant access to a user, a certain degree of trust is evaluated [12]. Delegation based trust model was proposed in multi-delegation model using trust management.

2.7 Bell and Lapadula (BLP)

Modified Mandatory Access Control (MAC) concept, which is quite similar to the Denning workflow model was proposed. It was observed that the BLP model is the extension to Denning and MAC on the basis of information flow. MAC in the BLP model provides security classification to object and subject [14]. Labeled assignment to object is called security classification, and those labels are considered for security clearance. BPL approach follows two rules: the first rule is simple security policy and the second rule is the property policy, which is also called the star policy. It's a kind of state machine used to enforce access control policy for military operations. It was developed by David Elliott Bell and Leonard J. LaPadula and is widely used to preserve confidentiality [15]. That's the main reason that it's widely used in government organizations and defense to protect secret and confidential data.

Fig. 2 describes the comparative analysis of ABAC and RBAC based on specific parameters. This classification is done using certain factors such as flexibility, security flaw, efficiency, authorization, and modification. It is observed from the DAC model analysis that this model is vulnerable to the Trojan horse attack. The only flexible model is RBAC. So it can be derived from the comparison table that the only rigid model is MAC, whereas DAC and RBAC are considered flexible.

Issues	RBAC	ABAC
Trend in 2018	Medium	High
Global Agreement	No	Yes
Flexibility	No	Yes
Easiness	Yes	No
Dynamicity	No	Yes
Authorization Decision	Locally	Globally
Granularity	Low	High
Manageability	Simple	Complex
Conviction	Locally	Globally
Confusing deputy	No	Yes
Changing privileges	Complex	Simple
Role explosion problem	Yes	No

Fig .2. Comparison of RBAC and ABAC using different parameters

Fig. 2 is showing us the identified issues in RBAC and ABAC models. A total of 12 issues were identified regarding two models.

3. Conclusions

The applications of access control in digital health-care systems play an important role in the present health-care industry, which can lead to automated data collection and verification processes, aggregated and correct data from different resources which can be immutable, tamper-resistant and deliver secured data, which leads in the reduction of probability of cyber-crime. Blockchain technology also supports distributed data, redundancy, and fault tolerance features for the digital system. In this proposed research, current challenges and problems in the digital health-care industry have been identified. In this paper, popular access models have been compared regarding their merits and demerits.

References

[1] A. Azaria, A. Ekblaw, T. Vieira, and A. Lippman, "Medrec: Using blockchain for medical data access and permission management," in *2016 2nd International Conference on Open and Big Data (OBD)*, 2016: IEEE, pp. 25-30.

[2] S. Azouvi, M. Maller, and S. Meiklejohn, "Egalitarian society or benevolent dictatorship: The state of cryptocurrency governance," in *International Conference on Financial Cryptography and Data Security*, 2018: Springer, pp. 127-143.

[3] A. Balu and K. Kuppusamy, "Cipher-text policy attribute based encryption with anonymous access policy," *arXiv preprint arXiv:1011.0527*, 2010.

[4] P. Beninger and M. A. Ibara, "Pharmacovigilance and biomedical informatics: a model for future development," *Clinical therapeutics*, vol. 38, no. 12, pp. 2514-2525, 2016.

[5] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *2007 IEEE symposium on security and privacy (SP'07)*, 2007: IEEE, pp. 321-334.

[6] M. Blaze *et al.*, "Dynamic trust management," *Computer*, vol. 42, no. 2, pp. 44-52, 2009.

[7] N. Broberg and D. Sands, "Paralocks: role-based information flow control and beyond," in *Proceedings of the 37th annual ACM SIGPLAN-SIGACT symposium on Principles of programming languages*, 2010, pp. 431-444.

[8] C. Cachin, "Architecture of the hyperledger blockchain fabric," in *Workshop on distributed cryptocurrencies and consensus ledgers*, 2016, vol. 310, no. 4.

[9] M. Chase, "Multi-authority attribute based encryption," in *Theory of cryptography conference*, 2007: Springer, pp. 515-534.

[10] R. Chinchani, A. Iyer, H. Q. Ngo, and S. Upadhyaya, "Towards a theory of insider threat assessment," in *2005 International Conference on Dependable Systems and Networks (DSN'05)*, 2005: IEEE, pp. 108-117.

[11] A. Das, M. M. Islam, and G. Sorwar, "Dynamic trust model for reliable transactions in multi-agent systems," in *13th International Conference on Advanced Communication Technology (ICACT2011)*, 2011: IEEE, pp. 1101-1106.

[12] N. Dimmock, J. Bacon, D. Ingram, and K. Moody, "Risk models for trust-based access control (TBAC)," in *International Conference on Trust Management*, 2005: Springer, pp. 364-371.

[13] J.-M. Do, Y.-J. Song, and N. Park, "Attribute based proxy re-encryption for data confidentiality in cloud computing environments," in *2011 First ACIS/JNU International Conference on Computers, Networks, Systems and Industrial Engineering*, 2011: IEEE, pp. 248-251.

[14] A. A. Donovan and B. W. Kernighan, *The Go programming language*. Addison-Wesley Professional, 2015.

[15] G. Zhao and D. W. Chadwick, "On the modeling of bell-lapadula security policies using RBAC," in *2008 IEEE 17th Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises*, 2008: IEEE, pp. 257-262.

A Permission Blockchain-based Chain of Custody at a First Glance

Seolah Jang[1] and Kyung-Hyune Rhee[2]
[1]Interdisciplinary Graduate Program of Artificial Intelligence on Computer,
[2]Dept. of IT Convergence and Application Engineering
Pukyong National University, Busan, South Korea
[e-mail: seolaang1020@pukyong.ac.kr, khrhee@pknu.ac.kr]
[2]Corresponding author

Abstract

The merits of blockchain technology have been heavily adopted in various applications, including digital forensics. The principal objective of implementing blockchain in digital forensics is to ensure the integrity of digital evidence. In the real world, there is a possibility that digital evidence is compromised not only during the confiscation but also in the post-imprisonment process. The evidence might be leveraged by multiple parties that take temporary ownership of it. Therefore, the immutable record of transactions is paramount in the digital forensics system. The characteristics of blockchain can be a plausible solution that can improve existing digital forensic operations. This process is well-known as a chain of custody. In this paper, we observe the essential points of implementing blockchain to address several challenges in the field of digital forensics. This study can be an initial consideration in the development of the digital forensics system.

***Keywords*:** Blockchain, chain of custody, digital forensics, digital evidence, Ethereum

1. Introduction

The current system of digital forensics has several challenges that have not been utterly addressed to this day, such as methods for maintaining the integrity of digital evidence. The shreds of evidence must be fully protected during the transferring, storing, and data collecting over the internet [1]. The digital evidence may be accessed and gathered by many different parties for investigation purposes. Therefore, the integrity of evidence is essential in digital forensics.

The merits of blockchain technology can cover up the drawbacks of the existing digital forensic system since it is tamper-proof without involving a third party to manage a transaction. Blockchain technology is also known as an immutable technology for storing information on distributed ledgers. These advantages can deliver various positive impacts, such as strengthening the integrity and reliability of transactions, securing transparency in a transaction, and reducing system costs [2]. In other words, the ledger which records transaction information is digitized and stored in a block.

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2020-2015-0-00403)supervised by the IITP(Institute for Information &communications Technology Planning &Evaluation) and partially was supported by This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2020-0-01797) supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation)

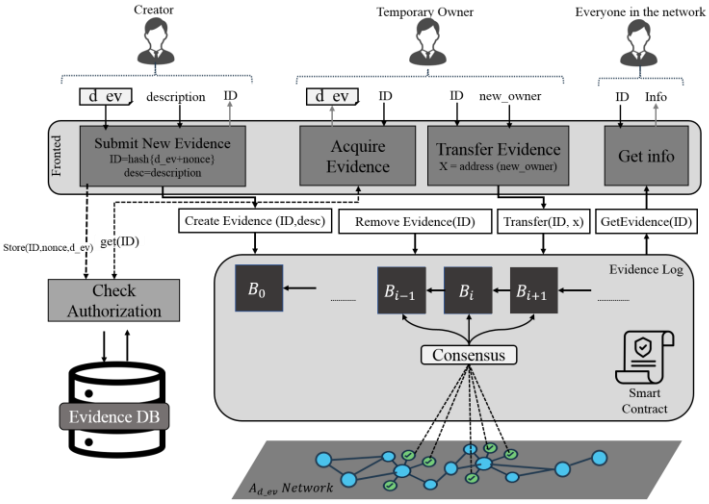


Fig. 1. Blockchain-based chain of custody mechanism

The integrity of evidence is the main issue in a legal court since digital data can be easily replicated and altered, unlike analog evidence such as bloodstains and fingerprints [3]. The prior scheme such as the blockchain-based chain of custody (*B-CoC*) [4] elaborates on the objective of combining blockchain technology in the digital forensics environment. However, several challenges must be addressed before blockchain is utterly adopted in the digital forensics systems, such as network topology, authentication mechanism, privacy, and security in the transaction [5].

In this paper, we observe several essential points related to blockchain forensics and chain of custody from the literature reviews. This study can be initial consideration in the development of the digital forensics system consolidate with blockchain technology.

2. Merging Blockchain Technology and Digital Forensics

The emergence of blockchain technology provides immutability in the investigation process of digital forensic. The terms of Digital forensic means a scientific investigation into the crime. Procedures and methods to establish and prove the facts of a particular act as a medium for digital devices in court. Digital forensics consists of three main requirements, which cannot be recognized as evidence without satisfying the following requirements:

1. Gathering evidence (recovery)
2. Analyzing evidence (reproduction)
3. Demonstrate integrity and write a report

The blocks in the blockchain networks are divided into body and header. More precisely, the transaction contents are stored in the body, while the hash codes that can prove the truth of every committed transaction are stored in the header. These blocks are propagated across the nodes in the blockchain network.

In regard to digital forensic evidence, every *B-CoC*'s activity carried out by the participant is recorded in its entirety into the body part. While the header part is stored the information of the previous blocks and Merkle roots information. Eventually, by combining these technologies, a reliable, secure, and traceable framework can be built incrementally.

3. Discussion and Outlook

Consensus selection. The consensus mechanism plays an important role in preserving efficient communication between parties in the *B-CoC* system. *Istanbul Byzantine Fault Tolerance* (IBFT) consensus protocol tolerates at most n faulty out of a total number of validators $v = 3n+1$. The latency of IBFT can be calculated as follows:

$$Ltc(x) = \frac{PPM_s + Pg_s + CX_s}{BW} \quad (1)$$

Where *Ltc(x)* is IBFT's latency, *PPMs* is the size of a pre-prepare message, *Pgs* is the size of the prepare message, *CXs* is the commit blocks' size, and *BW* is the network bandwidth. Since the clients are authorized in the *B-CoC* system, a consensus called proof-of-authority (PoA) is suitable to be adopted since it leverages the validators' identity to vote every incoming transaction as shown in **Fig. 2**.

By straightforwardly adopting a single consensus such as IBFT, proof-of-stake (PoS), proof-of-authority, might have many limitations in the real *B-CoC* system. Therefore, a consensus mechanism that can combine the merits of existing blockchain consensus is a necessity. This term is known as the *Hybrid* consensus.

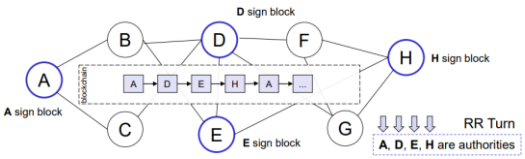


Fig. 2. Potential concensus for digital forensics [6]

Authentication process. The authentication scheme is framed by the parties' requests and server acknowledgments. The messages are exchanged in the B-CoC system that respects the authentication request. The acknowledgment can be either *SUCCESS* or *FAILURE*, where the *SUCCESS* response having no other content. While the *FAILURE* response commonly provides the following information:

- The information of authentication methods used by requester.
- The marker flag.

Platform selection. The gas usage and transaction fee may be a serious consideration in adopting the Ethereum platform in B-CoC. Moreover, every arbitrary change made in a smart contract is costly. Therefore, the platform such as Hyperledger Fabric might be a plausible solution to tackle the issues. It can also provide a pluggable consensus that can be used for any different system in the same environment.

Conclusively, the utilization of blockchain in digital forensic can tackle the weaknesses of the existing system, especially for management, transparency, and preserve an immutable record of transactions. A blockchain-based chain of custody for evidence management in digital forensics must meet several fundamental requirements as follows:

1. **Integrity**: the digital evidence must not be altered or modified during the process.
2. **Traceability**: the data flow of evidence must be traceable from the collection up to the temporarily used by multiple parties.
3. **Authentication**: the digital evidence only can be accessed by authorized parties. It is a part of an irrefutable sign as recognizable proof of parties' identity.
4. **Verifiability**: every activity must be provable from each entity which is also associated with their identity.
5. **Security**: blockchain is secure by design. However, the security mechanism in the authentication process is also essential. Moreover, changeovers of the evidence among parties should not be tampered with by malicious users.

4. Conclusions

We have presented several essential points in blockchain development in digital forensics systems with the main objective is to maintain the integrity of digital evidence from repeated use by different parties. By merging blockchain in the digital forensic system, several challenges are addressed accordingly. However, multiple factors still need to be considered thoroughly, such as consensus mechanism, authentication process, and types of platforms. Regardless of these matters, the characteristics of blockchain are paramount in managing the information flow of digital evidence. As future work, we are investigating how hybrid consensus can be utilized to improve system performance.

References

[1] A. Aminnezhad and A. Dehghantanha, "A Survey on Privacy Issues in Digital Forensics," *International Journal of Cyber-Security and Digital Forensics (IJCSDF)*. 2012.

[2] P. T. Duy, H. Do Hoang, D. T. Thu Hien, N. Ba Khanh, and V. H. Pham, "SDNLog-Foren: Ensuring the integrity and tamper resistance of log files for SDN forensics using blockchain," in *Proceedings - 2019 6th NAFOSTED Conference on Information and Computer Science, NICS 2019*, 2019.

[3] A. H. Lone and R. N. Mir, "Forensic-Chain: Ethereum Blockchain Based Digital Forensics," *Sci. Pract. Cyber Secur. J. 1(2)21-27*, 2017.

[4] S. Bonomi, M. Casini, and C. Ciccotelli, "B-CoC: A blockchain-based chain of custody for evidences management in digital forensics," in *OpenAccess Series in Informatics*, 2020.

[5] M. Cebe, E. Erdin, K. Akkaya, H. Aksu, and S. Uluagac, "Block4Forensic: An Integrated Lightweight Blockchain Framework for Forensics Applications of Connected Vehicles," *IEEE Commun. Mag.*, 2018.

[6] B. Rodrigues, E. Scheid, and R. Blum, "Blockchain and Smart Contracts – From Theory to Practice," pp. 1–31, 2019.

A Study on an Efficient Certificate-Based Aggregate Signature Scheme Providing Key Insulated in IIoT Environment

Yong-Woon Hwang and Im-Yeong Lee*
Department of Computer Science and Engineering, Soonchunhyang University
Asan-si, Chungcheongnam-do, 31538, Republic of Korea
[e-mail: hyw0123@sch.ac.kr, imylee@sch.ac.kr]
*Corresponding author: Im-Yeong Lee

Abstract

The Industrial Internet of Things (IIoT) is developing rapidly given recent developments in information and communication technology. Internet sensors in smart factories, smart grids, and medical environments are widely used to collect data in real time. However, if such data become exposed, data forgery/replay attacks are possible. Therefore, data encryption technology and digital signature technology are required. In this paper, we propose a new lightweight signature scheme based on a certificate-based aggregate signature that can generate and verify a signed message from a device in industrial Internet of Things environment. A gateway aggregates the signatures of messages generated by a group of devices, reducing the size of the entire signature and increasing the efficiency of operation. In addition, by providing key isulation, it is possible to solve the problem of key exposure due to physical attacks such as side channels. This can be applied to an environment where data is safely and efficiently collected in the emerging field of the Industrial Internet of Things.

***Keywords*:** IIoT; Certificate-Based signature; Aggregate signature; Key insulated;

1. Introduction

Currently, technological innovation is taking place in various fields with the advent of hyper-connected society. The Industrial IoT (IIoT) features sensors and peripheral devices facilitating communication and data processing in the fields of manufacturing, transportation, and energy; hundreds, thousands, or hundreds of thousands of terminal devices may be linked. If the network environment of IIoT sensors is well-configured, the Cyber Physical Systems (CPS) yields useful results [1].

However, if sensors transmit plaintext, data forgery/replay attacks are possible. In particular, millions of dollars worth of assets could be at risk if communications are not secured in large network systems such as the IIoT environment [2].

To solve this problem, a lightweight cryptography technology is required to provide data confidentiality, and a digital signature technology that provides reliability of the sensor device and integrity of the data generated by the sensor device in an IoT environment is very necessary. In the IoT environment, users can receive safe services only by providing integrity for data collected from devices. In addition, in an environment where large-scale data such as IIoT is collected, the concept of aggregation is important to efficiently collect data.

An early digital signature method used the public key infrastructure (PKI) encryption system. This was followed by signature and

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2020-2015-0-00403) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation)..

aggregate signature methods employing identity-based cryptography (IBC), identity-based signatures (IBS), identity-based aggregate signatures (IBAS), and certificateless cryptography (CLC). However, both the IBC and CLC methods pose key escrow and/or key distribution problems. To solve this problem, CBC-based signatures and aggregate signatures have been proposed, a studies are being conducted to meet a number of security requirements, such as forgery of signatures, data integrity, and non-repudiation [3-5].

In this paper, we analyze certificate-based aggregate signature(CB-AS) and propose a new lightweight CB-AS scheme in IIoT environment. In detail, this proposed method is an efficient Certificate-Based aggregate signature method that provides key isulation. The proposed scheme in this paper focuses on integrity rather than confidentiality of data. It will be applicable to a sensor device network environment that safely collects data continuously in real time in an IIoT environment that has expanded the general IoT environment.

2. Proposed Scheme

The eight phases include the KeyUpdate key insulation process and the seven phases of the CB-AS described in Section 2.3.1.

• Setup: The CA receives the security parameters and generates the public parameters, the master secret key, and the CA public key (step 1 of **Fig. 1**).

• KeyGen: Each sensor generates a public/private key pair using the public parameters and random values and requests a certificate from the CA. (step 2 of **Fig. 1**).

• CertGen: The CA generates a certificate $Cert_{\mathrm{ID}_i}$ corresponding to ID_i using ID_i per se, UPK_{ID_i}, a public parameter, and the master secret key of the sensor. The certificate is transmitted to the sensor. The CA creates a random value and sends it to the Helper (step 3 of in **Fig. 1**).

• KeyUpdate: The helper creates a temporary key hsK_{ID_i} for signing and passes it to the user, and the user creates a temporary key $Tsk_{ID_i,0}$ that can be signed. After that, the Helper creates an updated key $udsk_{ID_{i,t}}$ and sends it to the sensor, which then updates $Tsk_{ID_i,0}$ to $Tsk_{ID_i,t}$ for signature using the open parameters and updated keys received from the Helper (step 4 of **Fig. 1**).

• Sign: The sensor signs the message using the

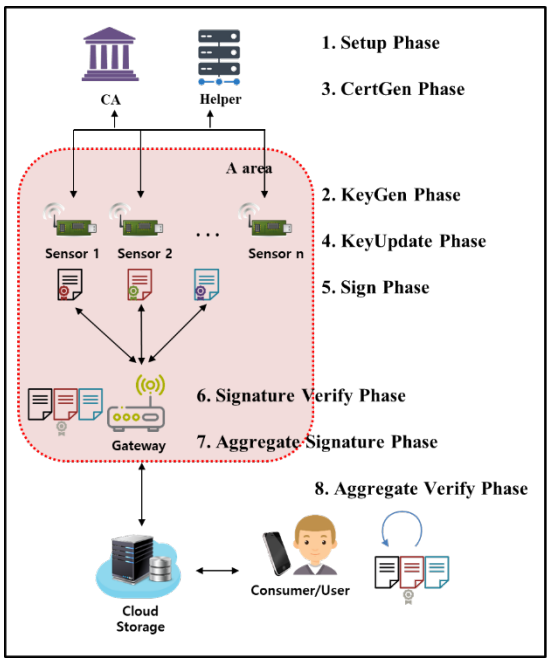


Fig. 1. Scenario of the proposed scheme.

ID_i, UPK_{ID_i} and certificate $Cert_{\mathrm{ID}_i}$ signature key $TSK_{ID_i,t}$ to send the message and signature to the gateway (step 5 of **Fig. 1**).

• Signature Verify: The gateway verifies the signature of the received message with ID_i, DPK_{ID_i} and certificate $Cert_{\mathrm{ID}_i}$ of each sensor device (step 6 of **Fig. 1**).

• Aggregate Signature: In this proposed method, the gateway serves as both a validator and aggregator. Messages and signatures $(m_1, \ldots, m_n, \sigma_1, \ldots, \sigma_n)$ collected from multiple sensor devices are verified with$(ID_1, \ldots, ID_i, UPK_{\mathrm{ID}_1}, \ldots, UPK_{\mathrm{ID}_i}$, $Cert_{\mathrm{ID}_1}, \ldots, Cert_{\mathrm{ID}_i})$ then aggregated through the gateway, and the aggregated data is transmitted to cloud storage. After aggregation, n signatures become one signature, reducing signature size. A verifier can check n messages in a single step (step 7 of **Fig. 1**).

• Aggregate Signature Verify: A user who wants to check the aggregated data can verify the integrity of the message collected from each sensor device by receiving the aggregated data $(m_1, \ldots, m_n, \sigma)$ from the cloud storage as a final verifier and performing verification with $(ID_1, \ldots, ID_i, UPK_{\mathrm{ID}_1}, \ldots, UPK_{\mathrm{ID}_i}, Cert_{\mathrm{ID}_1}, \ldots, Cert_{\mathrm{ID}_i})$(step 8 of **Fig. 1**).

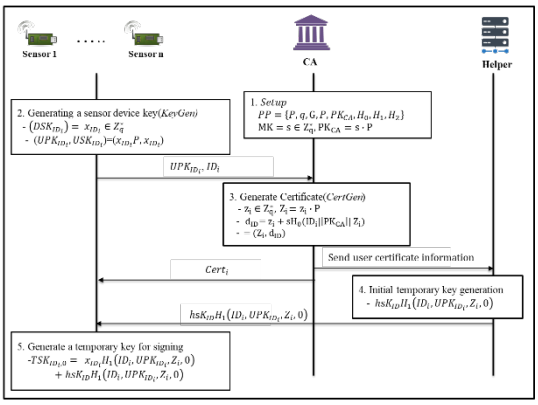


Fig. 2. Flow chart(Setup to CertGen)

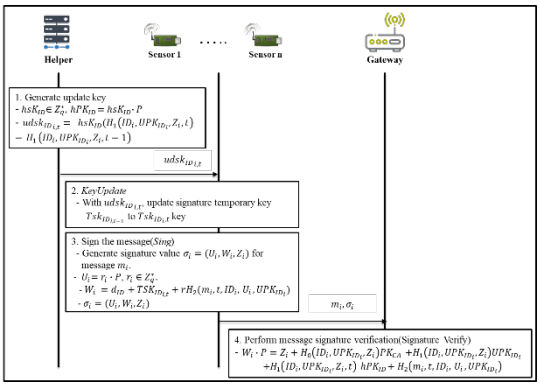


Fig. 3. Flow chart(KeyUpdate to Signature Verify)

Fig. 2~ Fig. 4 show the flow chart of the overall scenario.

3. Security Analysis of Proposed Scheme

• **Data Integrity and Reliability:** In this paper, the signature $\sigma_i = (U_i, W_i, Z_i)$ generated from each sensor device consists of a certificate value, a signature value including the user's private key, and a public key. The value to verify here is W_i. The value actually verified by the signature value σ_i to $W_i \cdot P$ operations is for $H_2(m_i,\ t, ID_i,\ U_i, UPK_{\mathrm{ID}_i})$, which can verify message integrity.

• **Unforgeability:** In the CB-PKC-based signature protocol, an attack on unforgeability can actually occur. The content of unforgeability is divided into an attack by an attacker capable of public key replacement attack and an attack by an attacker, a malicious CA with a lot of curiosity.

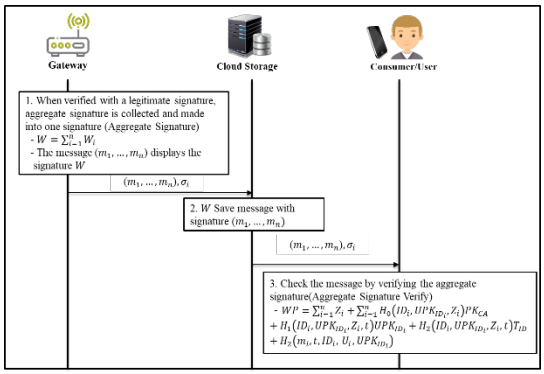


Fig. 4. Flow chart(Aggregate Signature to Aggregate Signature Verify)

has the ability to replace the public keys of other users with their own generated keys. Due to the safety of ECDLP, private keys corresponding to users' public keys cannot be generated, but only public keys can be replaced to bypass validation. Public key replacement attacks often occur in a certificateless method where there is no certificate that can authenticate the public key of a sensor device, which is a signer. However, even if it is based on a certificate, attacker can perform a public key replacement attack by creating a new private key/public key like the attack model. This is very difficult because it has to be performed within the polynomial time, and it is impossible because the public key is usually authenticated by a CA.

• **Key exposure due to side-channel attack:** In this proposed method, the sensor device initially registers through the CA, receives a certificate, and receives $hSK_{ID}H_1(ID_i, UPK_{ID_i}, Z_i, 0)$ from Hepler in the process of generating the key. And it creates a temporary key $Tsk_{ID_i,0}$ for signing. Then, $udsk_{ID_{i,t}}$ is received from the Hepler in the KeyUpdate Phase, and a key $Tsk_{ID_i,t}$ that can be signed for a period of t is generated and used when signing. That is, $TSK_{ID_i,t}$ includes the sensor device identifier ID_i, the private key/public key (USK_{ID_i}, UPK_{ID_i}), Hepler's private key hsK_{ID}, and some values of the certificate Z_i. At this time, even if $TSK_{ID_i,t}$ is exposed by a physical attack from an attacker, each sensor device can periodically update the

Table 1. Security and efficiency analysis of existing scheme and proposed scheme

Scheme / Items	Chen et al. Scheme [6]	Verma et al. Scheme [7]	Proposed Scheme
Type	CBAS	CBAS	CBKAS
Key exposure problem due to side channel attack	Vulnerable when key is exposed	Vulnerable when key is exposed	(Key update due to key isulated)
Forgery with Public Key Replacement	Safe (The signer's random value is included in the signature, so the attacker cannot know about the hashed message)	Not safe (Signature forgery is possible by attack model 1)	Safe
Forgery with CA Master Key		Not safe (Signature forgery is possible by security model 2)	Safe
Total Operation	$(3n+5)T_{sa}+(2n+5)T_{sm}+(n+1)h$	$(n+6)T_{sm}+nT_{sa}$	$(2n+8)T_{sm}+nT_{sa}$

T_{sm} : $Multiplication\ operation$; T_{sa} : $Addition\ operation$; T_{se}: $Exponentiation\ operation$; n: $Number\ of\ signature$;

signing key through the Helper and use it for signing.

4. Conclusions

The proposed scheme in this paper is for a certificate-aggregate signature affording key insulation for safe and efficient data collection in IIoT environments. As a feature of this proposed technique, the attacker cannot forge the signature, and the sensor continuously updates the signature key to solve the key exposure problem due to side channel attack. Even if a signature key becomes exposed, no forgery or camouflage attack is possible. In an environment such as IIoT, by collecting and aggregating messages signed by multiple sensor devices at the gateway, the total size of the signature can be reduced. The final verifier can verify the signatures of multiple messages with one verification, and this provides integrity, protection, and reliability for messages(**Table 1**).

For future research, it is thought that a study on signcryption is necessary to provide confidentiality as well as integrity of data communicated in IoT environment.

References

[1] Maede Zolanvari, Marcio A Teixeira, Lav Gupta, Khaled M Khan, and Raj Jain., “Machine learning-based network vulnerability analysis of industrial Internet of Things,” *IEEE Internet of Things Journal*, Vol 6, No. 4, pp. 6822-6834, 2019.

[2] Sadeghi, A. R., Wachsmann, C., and Waidner, M., “Security and privacy challenges in industrial Internet of Things,” *in Proc. 52nd ACM/EDAC/IEEE Design Autom. Conf. (DAC)*, pp. 1–6, Jun. 2015,

[3] Whitfield Diffie and Martin E., “Hellman. New directions in cryptography,” *IEEE Trans. Inf. Theory*, Vol 22, No. 6, pp. 644-654, 1976.

[4] Adi Shamir., “Identity-based cryptosystems and signature schemes,” *In Workshop on the theory and application of cryptographic techniques*, pp. 47–53. Springer, 1984.

[5] Al-Riyami, S. S., & Paterson, K. G., “Certificateless public key cryptography,” *In International conference on the theory and application of cryptology and information security*, pp. 452–473, Nov. 2003.

[6] Chen, Jian-Neng, Qun-Shan Chen, and Fu-Min Zou., "Certificate-based aggregate signature scheme without bilinear pairings," *Journal of Information Hiding and Multimedia Signal Processing* Vol 7, No.6, pp. 1330-1336, 2016.

[7] Girraj Kumar Verma, BB Singh, Neeraj Kumar, and Vinay Chamola., “CB-CAS: Certificate-based efficient signature scheme with compact aggregation for industrial Internet of Things environment,” *IEEE Internet of Things Journal*, Vol 7, No. 4, pp. 2563–2572, 2019.

[8] Xiong, Hu, et al. "Certificate-Based Parallel Key-Insulated Aggregate Signature Against Fully Chosen-Key Attacks for Industrial Internet of Things,", 2020

A Study on the Major Factors Affecting Government Policy in the Corona-19 Situation

Jung Chan Lee[1], Do Hoon Kim[2], and Bong Gyou Lee[3]*
Graduate School of Management of Technology, Yonsei University, Seoul, South Korea[1, 2, 3]
[e-mail: noared@nia.or.kr, dkim940627@gmail.com, bglee@yonsei.ac.kr]
*Corresponding author: Bong Gyou Lee

Abstract

The recent emergence of COVID-19, colloquially referred to as the Coronavirus, has radically altered the lives of citizens across the globe. The outbreak has incited a frenzy of sorts across a variety of industries, most visibly in the medical supplies sector. Similarly to the rest of Asia, the citizens of South Korea hurried to purchase masks as reports of a pandemic broke; a shortage of masks further fomenting panic across the nation. The Korean government-initiated countermeasures to quell the frenzy through the implementation of a 5-day rotation system for mask distribution and announced a series of regulations to take dominant control of mask supply. Despite the 5-day rotation system, public dissent persisted as the volatile nature of face mask supply continued to hamper comprehensive distribution and allocation to the public. Through the trials and tribulations, it has become clear that only through the collaboration between the South Korean government and the private sector, the government was able to develop platforms such as the public mask inventory alert service through the utilization of public data accumulated by contractors in the private sector. Since it's conception and release, the aforementioned public mask inventory platform has done volumes in terms of easing the anxiety of the South Korean people. Learning from the face-mask crisis, we were able to derive that effective data sharing is integral to the continued growth of public health programs and its accessibility. Through this study, I would like to studied the causal structure in which communication factors affect government trust and data open policies affect acceptability in the process of promoting public mask information open policy.

Opening of public mask information for COVID-19 prevention policy. It is expected that it will serve as a major communication factor in the policy-making process, including the degree of fairness in policy decisions, the degree to which stakeholders are aware of the opportunity to participate, and the degree of transparency in policy making. It is expected that trust in the government and trust in the public mask information opening policy decision itself will have a positive effect, and trust in the government and policy will have an effect on the acceptance of the public mask information opening policy decision.

***Keywords*:** COVID-19, crisis management strategy, corporate, pandemic, public data, ICT

1. Introduction

The Korean Government is actively using ICT(information communication technology) and public data to managing face mask inventory information to control confirmed COVID-19 patients and spreading of COVID-19. Most importantly, providing face mask inventory information to frighten people is a successful action to prevent the spreading of COVID-19.

Private developers developed various face mask data notification web or mobile application services using the government's open public data. Furthermore, various map-based mask app services were developed simultaneously by

opening the location information of pharmacies provided by the government by converting and converging GPS-based coordinates information, dramatically reducing the effort and confusion of wandering around by searching for masks. These ICT services have become representative examples of Civic Hacking, in which civil society is the main body of problem-solving. Our research investigates how the government's urgent policy of opening public mask information as a precautionary measure of Corona-19 has affected the people's trust in the government and its policies.

In this study, we would like to focus on the fairness and transparency of the policy-making process, which are the key factors to consider in the policy-making process, and the recognition of private civil organizations to provide opportunities for participation in decision-making of relevant data-opening policies. Stakeholders' participation in all of the government's policymaking and enforcement process is impossible. Stakeholders' participation in all of the government's policymaking and enforcement process is impossible. On urgent and disastrous issues such as the COVID-19 crisis, the involvement of many stakeholders hurts the part of urgent action in policy decision-making and enforcement.

In recent years, the Korean government put extra effort into gathering public opinions on significant policy decisions. Still, not many studies have been conducted on the participation or consideration recognized by the public and stakeholders in policymaking.

Through this study, I would like to study the causal structure in which communication factors affect government trust and data open policy affects acceptability in the process of promoting public mask information opening policy. At the time of preparation for post-Covid-19, we would like to empirically examine what factors affect government policy in the situation of national crisis.

2. Theoretical Background

1) Crisis Management and Policy Acceptance

Most of the prior studies of policy acceptability are conducted to identify policy acceptability in the face of conflict. The environment of conflict arose from social problems such as conflicts between government-citizens, citizens-citizens, and citizens, and the main topics are policy acceptance in the face of disputes over the location of hate facilities (nuclear power plants, radioactive waste), and policy acceptance in the environment of conflict between the government and citizens over international relations policies such as FTA.[1]

Choi (2018) analyzed the factors determining the acceptability of policy implementation between the general public and residents of conflict-experienced areas in the policies for installing power-related facilities (such as nuclear power plants, waste dumps, power transmission towers, etc.) and the guidelines for accepting/installing military facilities (such as THAAD deployment, naval bases). According to the analysis, if the policy target can grasp information and financial ability, it positively affects policy acceptability. Also, When the policy implementer's tried to comply with laws and guidelines to ensure resident's participation, policy acceptability tended to change more positively.[2]

Various studies on policy acceptability have been carried out steadily in the conflicts of social organizations, research on the acceptability of policies pursued in disaster management situations such as the COVID-19 crisis, the MERS outbreak, earthquakes, and floods is insufficient, and research on citizens' policy acceptability in the process of accepting policies during disastrous conditions is also hard to find.

Kim(2016) studied that the responsibilities of the government and bureaucrats play an essential role in policy acceptability in dangerous situations by comparing the spread of MERS with the case of the Ferry Sewol disaster.[3]

Lee(2016)investigated the relationship between the government and citizens by analyzing the media data in the process of responding to MERS in Seoul and by analyzing the speed, consistency, openness, control tower, and use of the crisis manual. The Seoul Metropolitan Government initially failed to lead the crisis, but later suggested that all factors in the response process evaluated to have increased the citizens' policy acceptability. [4]

However, previous research papers failed to analyze the public opinion of citizens shown on the media and SNS according to the time flow of the critical situation and was unable to suggest what variables influence the policy acceptability of citizens in the crisis management condition.

2) 4th Industrial Revolution and the Change of Policy Acceptance Environment

With the proliferation of intelligent information technology, explained by the convergence of ICBM (IoT, Cloud, Big Data, AI, Mobile), communication channels on policy acceptability have diversified. People are simultaneously accepting and criticizing governments' policies and pouring out vast amounts of information through online discussion rooms, SNS hashtags, and one-person media channels. Thus artificial intelligence(AI) technology has also been developed to check news generated through enormous amounts of big data. [5]

In the past, the government unilaterally communicated its policies to citizens. Due to the rapid development of ICT, it has changed to an environment where the government and citizens express their opinions in both directions, such as 'Blue House Online Citizens' Petition.' The development of the information and communication environment also creates a drastic change in the policy acceptability of citizens.

"The government's public data opening policy in the early stages of the Corona-19", development of mask inventory information app for private developers using public data, using this information, this is an example of the changing policy demand environment after the 4th Industrial Revolution, such as the development of mask information apps by private developers and the joint development of 'mask maps' in accordance with the 'Gwanghwamun 1st Avenue' innovation proposal.[6]

However, prior research has not yet been conducted to analyze this series of data opening policies. Therefore, in this paper, we would like to hypothesize and analyze the impact of technology related to the 4th Industrial Revolution on the policy acceptability of citizens.

3) policy acceptability

The government operates a channel for communication between the government and the people through 'Gwanghwamun 1st Avenue' to reflect it in significant policies through communication to the public. The rate reflected in the procedure is also increasing. The method of communication between the majority of the people and the government is an essential element in improving policy acceptability.

George & Sharkansky(1978)[7]and Chae(2017) presents Policy content (faithfulness, consistency), policy implementer (performance, interaction will, trust), policy targets (protection experience, personal characteristics), policy environment (media impact, relevant local situation), policy means (regulatory and economic compensation, education and publicity, procedural factors). [8]

Cho et al. (2019) said policy acceptance factors include policy participation, governance, trust through the establishment of communication systems, information provision, policy promotion, and learning. [9]

Jeon et al. (2016) set government trust, policy satisfaction, policy acceptance, policy expectations, policy performance, and expectation disagreements as factors. [10]

This study was conducted by referring to the research design of Jung et al. (2017), which studied the relationship between policy decisions, fairness in policy decisions, awareness of stakeholders' awareness of opportunities for participation, and the degree of transparency in policy decisions as independent variable.

3. Research Model & Methodology

1) Hypothesis

This study synthesizes the theoretical background of previous studies and covers the fairness of policy decisions to open public mask information for the COVID-19 prevention policy, the degree of awareness of the opportunities for participation by stakeholders, and the degree of understanding of transparency in policy-making in the process It is defined as the primary communication element which structured the theoretical relationship that they influence the government's trust and trust in the public mask information open policy decision itself, and policy trust influence the acceptance of the opening public mask information decision. Therefore, we assumed that the recognition of the benefits of public mask application or web service will have a moderate or interactive effect on policy acceptance. This research's hypotheses are as follows,

(**H1**) The perception of fairness in determining the policy of opening public mask information will have a positive (+) impact on trust in government.

(**H2**) The perception of fairness in determining the policy of opening public mask information will have a positive (+) impact on trust in policy.
(**H3**) The perception of stakeholders can participate in the policy process of opening public mask information will have positive (+) impact on trust in government.
(**H4**) The perception of stakeholders' can Participate in the policy process of opening public mask information will have a positive (+) impact on trust in policy.
(**H5**) The perception of transparency awareness of open public mask information will have a positive impact on trust in government.
(**H6**) The perception of transparency awareness of open public mask information will have a positive impact on trust in policy.
(**H7**) The governments' confidence in the opening public mask information policy will have a positive (+) impact on policy acceptance.
(**H8**) The degree of trust in the opening public mask information policy will have a positive (+) impact on policy acceptance.

2) Data Analysis

The sample population of this study collected data through the online survey method by setting up the general citizens using public mask-related data services as Unit of Analysis. The survey was collected for two weeks from Oct. 5 to Oct. 16, 2020, with 250 responses being received and 245 responses were used. We excluded five insincere answers. The collected data were verified using SPSS 18.0 for factor and reliability analysis, and SmartPLS 2.0 was used to verify the hypothesis in this study.

A factor analysis was conducted for each variable to verify the collected survey data's validity and reliability. The factor analysis was analyzed using the equamax method with Kaiser Normalization. The reliability analysis for each factor identified by the factor analysis determined that the Cronbach's alpha value was 0.7 or higher. [11]

The factor analysis results confirmed that the composite reliability and the extracted average variance were more than 0.8 and 0.5, confirming the research model's convergent validity in this study. And, in this study, the communality values investigated are all 0.5 or higher, satisfying the measurement model's suitability.

Partial Least Square(PLS) was used to verify the hypothesis for this study, and SmartPLS 2.0 was used as a tool for analysis. By structural model analysis, the coefficient of determination of path and endogenous variables were derived, and the total of 245 survey results was bootstrapped to 500 to verify the significance of the hypothesis. The verification of the structural models is as follows. [12]

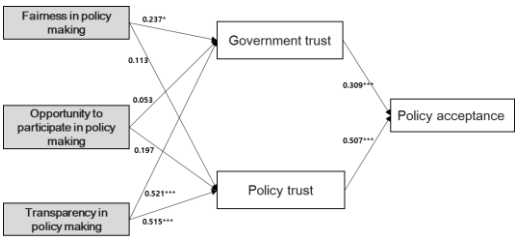


Fig. 1. Data analysis

4. Results

Opening of public mask information for COVID-19 prevention policy It is expected that it will serve as a major communication factor in the policy-making process, including the degree of fairness in policy decisions, the degree to which stakeholders are aware of the opportunity to participate, and the degree of transparency in policy making. It is expected that trust in the government and trust in the public mask information opening policy decision itself will have a positive effect, and trust in the government and policy will have an effect on the acceptance of the public mask information opening policy decision.

5. Conclusion

This study intends to identify causality on government trust and policy acceptability by opening public mask information. Furthermore, we wanted to empirically examine what factors affect government policies when Corona-19 is spreading, and similar national disasters occur.

The main findings that we have identified in this study are as follows. First, the recognition of fairness in the process of determining the opening public mask information policy was found to have a positive (+) impact on the government's trust. However, it was rejected in the provision of policy trust and opportunities for the participation of stakeholders. Meanwhile, the transparency recognition of the opening public

mask information policy had a positive (+) impact on the government's trust and the policy's trust. Due to the fairness and transparency of government policies, the process of trusting the government and accepting policies has shown significant results.

The results of this study can provide policy implications for future policy acceptability. First, it is essential to provide transparent information related to the fairness and policy of policy decisions during the government's new policymaking. Our study claims that the public's awareness of wanting quick and accurate information in the COVID-19 pandemics includes fair and transparent policy decisions in the policymaking process.

In addition, various stakeholders' participation is an important factor in the policymaking process. However, stakeholder's participation did not have a significant impact on our study. It can be said that in a global disaster such as the COVID-19 pandemic, the government's transparency, and fair, rapid decision-making is more important. Results of these studies may infer that government-led quick and fair decision-making is important in global disaster and that the role of the central control should be further strengthened.

Reference

[1] Chae.JongHyun(2015), A Study on Policy Design for Policy Acceptance Improvement and Conflict Prevention, KIPA

[2] Jo.GongJang(2019), A plan to improve social acceptance of environmental policy: focusing on environmental culture, KEI

[3] Kim,Byoungseop(2016) Risk Society and Bureaucratic Responsibility, Korean Society and Public Administration Research

[4] Lee,Inwon(2016) A Study on Crisis Management Communication from the Perspective of Policy PR: Focusing on the MERS Response Process in Seoul

[5] Son,Sangyoung(2019) Research on national strategy for securing social acceptance of the 4th industrial revolution, KISDI

[6] DongA-ilbo(2020.7.14), Given time is a week, develop a COVID-19 mask app?

[7] Edwards, George C., III, Sharkansky, Ira. (1987). The Policy Predicament. San Francisco: W.H. Freeman, Publication date

[8] Chae.JongHyun(2017), A Study on Enhancing Policy Acceptance and Social Integration through Promotion of Public Debate Procedures, KIPA

[9] Lee,Jongyeop(1997) A Study on the Policy Acceptance in the Decision and Implementation Process of Location Policy-Focused on the Cases of Location Selection for Hate Facilities, Chungnam National University

[10] Jeon,Younghwan(2016), The Influences of Risk Perception and Trust in Government on Nuclear Energy Policy Compliance and Satisfaction: Applying the Expectancy Disconfirmation Mode, KJPAE, 2016, vol.26, no.3, pp. 85-110

[11] Hair, J. F. 1998. Multivariate data analysis, (Upper Saddle River, N.J. : Prentice Hall: Upper Saddle River, N.J.

[12] Fornell, C., and Larcker, D. 1981. "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," Journal of Marketing Research:1), p 39.

NepaliKisan: A Web based Agriculture Management Information System (A-MIS)

Deepanjal Shrestha[1], Tan Wenan[2]* Neesha Rajkarnikar[3], Deepmala Shrestha[4] and Seung Ryul Jeong[5]*

[12]School of Computer Science and Technology
Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
School of Energy Science and Engineering, Nanjing Tech University, China
[14]School of Business, Pokhara University, Nepal
[5]Graduate School of Business IT, Kookmin University
Seoul, South Korea
[e-mail: deepanjal@pusob.edu.np, wtan@foxmail.com, neesha_rk@hotmail.com, f13deepmalas@iima.ac.in, srjeong@kookmin.ac.kr]
*Corresponding author : Tan Wenan, Seung Ryul Jeong

Abstract

Agriculture is the biggest industry of Nepal and faces various challenges at social, economic and governance level. The gap between producer market, consumer market, lack of proper information channels, barriers to trade and transport is attributing for workforce migration, ageing and declining of this industry. "Nepali-Kisan" is web based management information system designed for resolving the gap between producer and consumer market, overcome the problem of information gap and utilize the current trade and transport model to empower farmer in the form of small-scale enterprise industry. This work proposes an integrated digital model to facilitate business and governance processes between farmers, industries, government agencies and transportation agencies to facilitate their business. An exploratory approach with system analysis and design principles is applied to build Nepali-Kisan web information management system. Actors, business processes, governance regulations and dependencies are identified to build a conceptual model, transport management model and integrated web based information management model for the Nepalese agriculture industry. A verification and validation matrix is developed to test the model at various levels of integration. The system has vital importance for farmers, small scale enterprises and consumers as its provides an open platform for open and fair market increasing satisfaction at all levels. The government bodies, NGOS, INGOS, GOs working in this sector can obtain precise and correct information contributing in the national governance and agri-business management.

***Keywords*:** Agriculture MIS, Supply Chain, Agriculture product, Information access, Nepal, Farmers

1. Introduction

Nepal is an agricultural country where more than 66% of the total population is involved in agriculture as their means of livelihood [1]. The percentage contribution to GDP by agriculture is 27% which indicates that Nepal has a prodigious scope in uplifting the national economy through agriculture [1]. Despite of this fact, Nepal imported Rs 220 billion worth agricultural goods in the last fiscal year [2]. Nepalese agricultural industry faces challenges from market place, climate change, land destruction and workforce migration leading to ageing and declining of this industry [1]. Further, the government prioritizes this sector an allocates special budget, schemes

and policies but has provided to significant improvement in the declining status of this industry [3]. The government has proven inefficient in defining standard prices and quality parameters for agriculture product and service and is unable to reach to every section of farmer due to difficult terrain and lack of information and communication infrastructure [4]. There is a decreasing interest of modern Nepali youth in this industry as the gain in income is less compared to working in other sectors like foreign land or get employed in some other business [5]. This has ultimately hit the current and future perspective of this industry. It is seen that this industry is badly influenced by many actors who work on the supply and demand side of agricultural industry and decide every aspect of this business. Collecting crops and vegetables in cheap price from farmers and selling it at high price at consumer side has greatly affected this industry in Nepal [5]. The predominance of intermediaries has created a dissatisfaction both at producer and consumer side with government and regulatory bodies unable to check and control this issue.

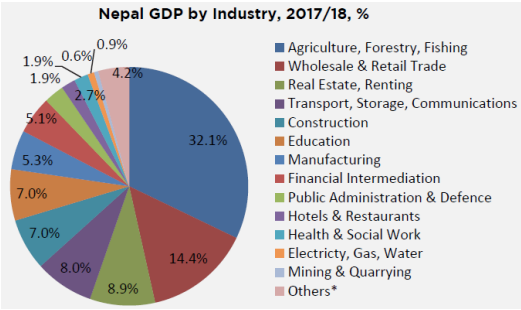


Fig. 1. Nepal GDP by Industry, Source [14]

The intermediaries bargain with farmers and are forced to sell what they have produced in their own villages or bound to sell their product at very low price to private money lenders, brokers or traders which make them gain very less or no profit from their cultivation [6] [7]. On the other hand, the same supply chain brokers keep the consumer away from getting cheap and good quality agricultural domestic products in time [1] [4] [6]. The poor communication and market reach of farmers make them less aware of the current market price of their products and they are bound to sell them at prices set by intermediaries and business men.

The role of intermediaries, the influence of the Indian and Chinese market, lack of a master plan, poor storage; processing, transportation facilities and weak market intelligence are the major factors that are hitting the Nepalese agriculture industry at present context [9] [10]. Further the lack of information on agriculture products, disease in plants, climatic conditions, remedies and cure is inaccessible to a large mass of farmers [11]. There is no active digital implementation to look into this factor and the vital information regarding the agriculture is limited to some places and out of reach for the majority of the farmers. Despite continuous efforts and approaches made by the country to combat these hurdles there is little or no achievement. There are numerous setbacks in this field and that need to be displaced. The government has been pushing its efforts from different angles by educating farmers, bringing awareness programs, checking markets but still the output is very low and disheartening [12].

It has been realized that the application of technology and shifting of technologies from the innovators to the farmer's level will be the most important achievement to overcome this unfair system [12]. Various GOs, NGOs, INGOs are working in this direction and are providing necessary inputs. Besides these, today different applications (mobile and web app) are working as the pocket guideline of the farmers in various perspectives, like some of the mobile applications are designed to help the farmers in crop management, information about weather related data, the opportunities in the agricultural field, suggestions from the experts, answers of the queries of the farmers etc. [13] [14] [15].

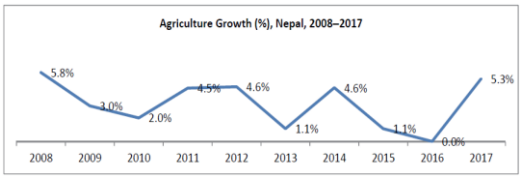


Fig. 2. Growth in Agriculture Industry, Source [14]

There is a need to answer the holistic problem of Nepalese agriculture integrating solutions at business, technology and governance level. "Nepalikisan" is undertaken to overcome all the above listed problems. The system links up farmers to both retailers and consumers creating an open platform for all stakeholders of agriculture business helping them to conduct business and information exchange.

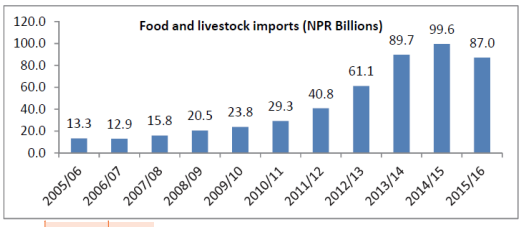


Fig. 3.[U1] Food and livestock import of Nepal. Source [14]

2. Objective

The research has the following set of objectives to fulfill which are mentioned below:

1. The system must be able to bridge the agricultural gap through knowledge management through researchers, extension agents and farmers enhancing agricultural production
2. The system must be able to market agriculture products and services through open platform with marketing and trade in a variety of ways
3. The system must be able to standardize product quality with traceability to help in delivering more efficient and reliable data to comply with international standard.
4. The system must support financial support and risk management and increase access to the financial services for rural communities helping to secure savings find affordable insurance and tools to better manage risk.
5. The system must be able to support capacity enhancement and provide empowerment opportunities, educating local communities to provide new business opportunities enhancing livelihood.

3. Research Framework

The study employs software engineering tools and techniques to analyses and design the system and domain analysis is performed considering the technical literature, existing design and customer requirements. The design model produces use cases, class model, information model and system model as shown in **Fig. 4**.

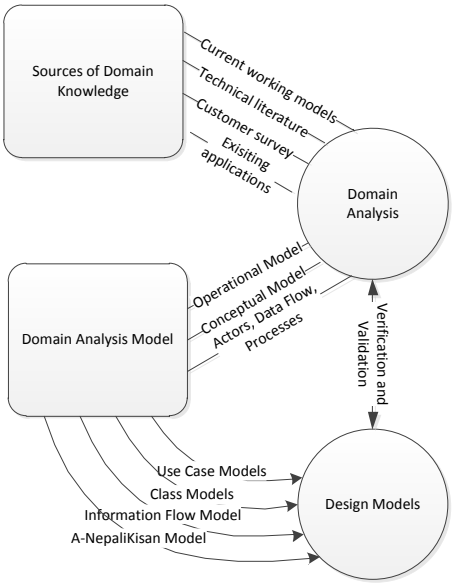


Fig. 4. Modified framework of research. Source [16]

4. System Analysis and Design

System analysis is a systematic approach of investigating the problem domain to identify the elements of information system. An investigation of the current system in practice is taken into consideration to highlight the basic issues of problem domain. The system analysis is done to identify, the artifacts, concepts, scenarios and functional processes of the system.

4.1 Operational Model

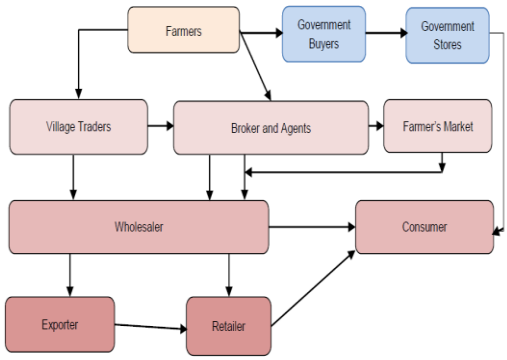


Fig. 5. The Operational model of agricultural business in Nepal

This operational model explains the current system in practice where a farmer is connected to three major players that include Villager trader, Broker agents and Farmers market. These three players are the first layer of contact for the

farmer and the whole bargaining of the products is decided by these players. The villager and the commission agent are further connected to the wholesaler who forms the second layer and arranges for the selling of products and services through the retailers. The first layer is the first responsible category that decides the price followed by second layer where the whole seller is another player who plays his role in increasing the price of the agri-products. The retailers are the final players who play a strategic role in finally deciding the price before it reaches the consumers as shown in **Fig. 5**.

4.2 Identifying problems, concepts and artifacts

Table 1. General problems of agriculture industry. Ref []

S.N	Problems
1	Poor and inadequate access to agricultural input and supply
2	Declining productivity, low yields and poor management
3	Poor irrigation and water reach
4	Limited financing availability and poor incentives
5	Lack of free trade, poor and dependent access to markets, Role of intermediaries
6	Labor shortage, ageing problem and labor migration.
	Poor transport and distribution channels

Table 2. ICT and digital solutions for Agriculuture industry. Ref []

S.N	ICT and Digital Components
1	Open and Integrated Agriculture Business Platform must be designed and implemented connecting all stakeholders.
2	Introduce mobile applications to provide information on the weather, market information, prices, and crops
2	Leverage digital technologies for support and information on agriculture tools, technology and machinery.
3	Virtual, online and remote education to impart technical knowledge and best practices in agriculture.
4	Improve productivity through the use of technology like satellites, drones, and soil sensors for better information on weather, soil and environmental changes.
5	Using IoT and sensor technology to ensure optimization of resources and real time monitoring
6	Improvement of irrigation systems through smart management to minimize water loss and ensure higher irrigation efficiency.
7	Online selling, advertising and facilitating digital payments to farmers, government liaisons, intermediaries and consumers.
8	Crowdsourcing, Big data analysis, social network connections for creating knowledge base and consumer market with better analysis and prediction systems.
9	Use of logistic solutions and sensors to track trucks and obtain location updates for transport of goods
10	Provide matching platforms to help grade the quality of produce
11	Deploy traceability and tracking systems and RFID for smart packaging
12	Automation for ploughing, leveling, winnowing, harvesting, spraying, and irrigating reducing the need for manual labor

The sources of domain knowledge are further investigated to identify the general problems in the agriculture sector of Nepal as shown in table 1. An ICT and digital solution is traced for the general problems identified and mapped as a proposal and solution for the current and future needs as shown in table 2. Tabel 3 represents the actors, their roles and responsibilities in a system under investigation. Altogether 18 actors are identified that consitute an important component of the business. These actors are detailed with their role in the current system and their responsibilities they are bestowed with. This table along with other two tables and operational model is used for further analysis and design of the NepaliKisan Information Management System.

Table 3. Representing Actor, their roles and responsibilities

S.N	Actor	Roles and responsibilities
1	Farmer	Produces, Sells, Updates knowledge of agriculture, Reports problems
2	Ministry of Agriculture, National	Allocate budget, Provides financial support, Regulates, monitors, manages and provides information, Takes Complaints, Provide Solutions, Updates technology in agriculture, Orients international practices, Organize workshops and conferences
3	Ministry of Agriculture State	Allocates budget, Provides financial support, Regulates, monitors, manages and provides information, Takes Complaints, Provide Solutions, Updates technology in agriculture, Orients international practices, Organize workshops and conferences
4	District Agriculture Office	Regulates, monitors, manages and provides information, facilitate government schemes and privileges, Takes Complaints, Forwards Complaints, Provide Solutions
5	Rural Agriculture Office	Regulates, monitors, manages and provides information, facilitate government schemes and privileges, Takes Complaints
6	Agricultural Extensions	Make research, motivate agriculture systems, Orient and teach farmers, Takes Complaints, Forwards Complaints, Provide Solutions
7	NGOS, INGOS, GOS	Research and work with agriculture sector and community of a particular area. Updates technology in agriculture, Orients farmers, Organize workshops and conferences, Provide finance and support

8	Farmer Selling Extensions	Buys, Updates prices, provides support
9	Retailers	Buys and sells agriculture products
10	Financial Institutions	Provide loan for agriculture on subsidy
11	Consumers	Buys products, evaluates product, reports
12	Agriculture Markets	Buys, sells and Manages agriculture products
13	Wholesalers	Buy agriculture products, sells to retailers, consumers
14	Intermediate Buyers	Buy agriculture products, sells to wholesalers, retailers, consumers
15	International Buyers	Buys products, Validate quality, Mechanism international market
16	Government buying extensions	Buys agricultural products, store and sell at subsidized rates, Takes Complaints, Forwards Complaints, Provide Solutions
17	Transportation Industry	Carries agriculture products from farmer site to markets
18	International donors	Work with Government agencies, provide grant in agriculture sector, work with community to promote and develop agriculture. Updates international practice and Orient farmers, Organize workshops and conferences

4.3 Use cases models

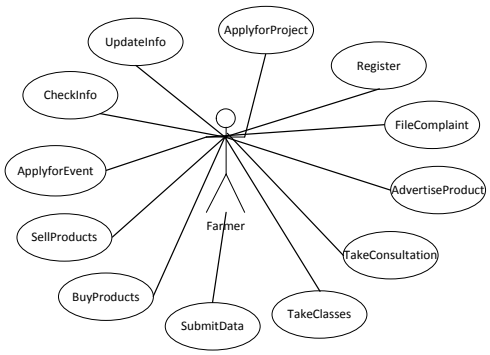


Fig. 6. The Use case model with farmer and its related processes

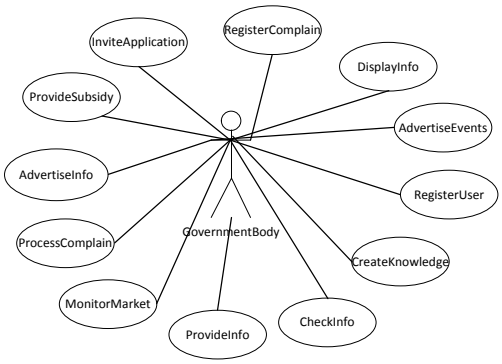


Fig. 7. The Use case model with Government body and its related processes

Use cases are very important part of system analysis and design. Here we can see that major actors with their related processes are depicted [16] . **Fig. 6** represents farmer with its related processes which include RegiterUser, SellProducts, ApplforProject, FileComplain, etc. Similary, **Fig. 7** represents tha Government body as an actor with its related use cases RegisterUser, InviteApplication, DisplayInfo, OrganizeEvent, etc. Figure 8 shows Business agent which forms another important component of the business. This actor uses TrackOrder, SellProduct, SubmitInfo, Register, etc. as the major use cases of the system which allow him to interact and complete a particular functional process.

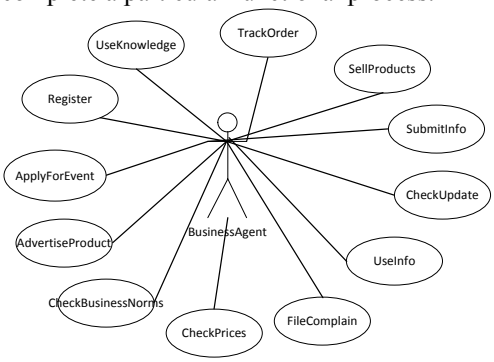


Fig. 8. The Use case model with business agent and its related processes

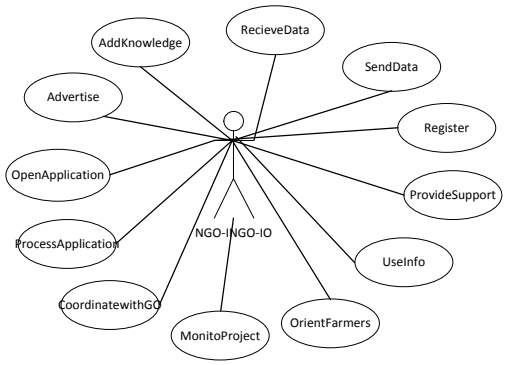


Fig. 9. The Use case model with INGO-GO-IO and its related processes

The (INGO) International Non Governmental Organizations, (GO) Governmental Organiztions and (IO) International Organization form very important part of agriculutral development and research in Nepal. These organization work in close proximity with the governmental bodies and help the farmers and communities to develop itself and improve their life standards. These

agencies provide funds, technical assistance and make research on local and customized way to improve the agricultural sector of Nepal.

4.4 Conceptual Class Model

Concepts are building blocks of the system which later are extended to from Class diagram that represent the dynamic modelling of the system [16]. Use case, conceptual diagram and sequence diagrams can be later mapped to collaboration diagram that depicts the dynamic behavior of the system. Figure 10 represents the conceptual model of the system that are linked with one another in a generalized relationship. The identified concepts are the most important and major aspects of the system that participate in the system to accomplish a particular functional process of a system. Altogether 16 major concepts are identified in the system that include farmers, governmental bodies, INGOS, GOs, NGOs, Wholesalers, International buyers, Consumers, business agents as shown.

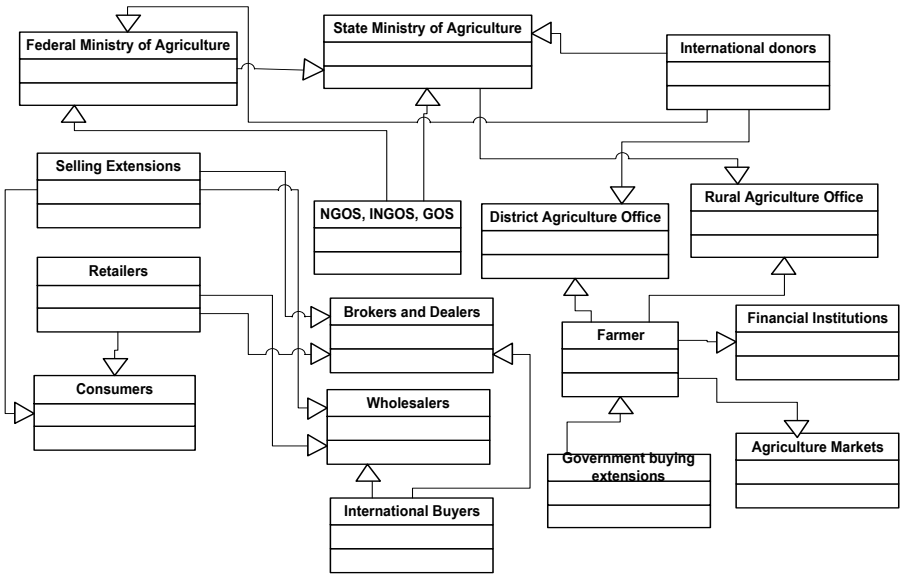


Fig. 10. The Conceptual Model representing concepts in a generalized relationship

4.5 NepaliKisan Information Management System.

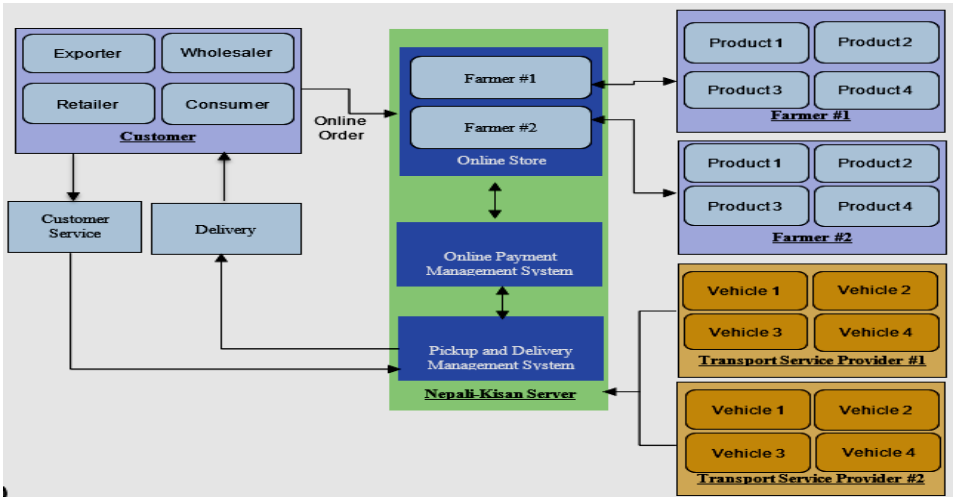


Fig. 11. NepaliKisan Information System Management Model

This model explains the proposed system where a farmer is directly connected to all the major players that include exporter retailer, wholesaler and consumer connected. The farmer in this model has direct contact with these players and can play a vital role in deciding the value of his products and services. This model is inspired by ICT and opens business systems where all the consumers and producers gather in one place and can bargain for their product and services. This model promotes open market but also challenges non-standard prices and quality of assurance. It is highly dependent on Trust and Value Creation and works virtually. If it meets its prescribed parameters for quality and robust network, it is one of the best systems for implementations. The proposed system is taken into consideration to overcome the problems faced by Nepali farmers and promote agriculture business in Nepal as shown in **Fig. 11**. This information system benefits smallholder farmers to commercial farmers by raising their bargaining power and increasing their incomes (by enabling them to get information related to their inventories and negotiate for fairer prices), reducing price dispersion (variation which creates market inefficiencies) and by reducing or stabilizing the possible risks.

5. Conclusion

The scope of Information and communication technology in agriculture (ICT in agriculture), e-agriculture is huge for enhancing agricultural development. This study is the outcome of literature in agriculture marketing systems in Nepal and has been conducted using published materials, web resources and for Nepal an interactions and interviews with the agriculture supply chain players and stakeholders. Hence, it is concluded that integrating famers with technical assistance can be a prudent plan to enhance agriculture productivity, product quality/quantity and improve supply chain management which not only foster national economy but also enable good governance and management in this sector.

References

[1] International Labour Office, ""EIGHT WAYS TO GROW NEPAL'S AGRICULTURAL SECTOR, ISBN: 978-92-2-133721-8 (print), 978-92-2-133722-5 (web pdf), Printed in Switzerland, Copyright © International Labour Organization 2019, https://www.ilo.org/wcmsp5/groups/public/---ed_emp/documents/publication/wcms_713334.pdf

[2] Sangam Prasain, Alarm bells are ringing as agro imports bill reaches Rs220 billion, The Kathmandu Post, 28, 2019, https://kathmandupost.com/money/2019/07/28/alarm-bells-are-ringing-as-agro-imports-bill-reaches-rs220-billion.

[3] Chaudhary, Deepak, AGRICULTURAL POLICIES AND RURAL DEVELOPMENTIN NEPAL: AN OVERVIEW, Nepal Journal of Development Studies, Nepal Journal of Development Studies (Year 1st Issue 2nd, 2018 November), https://www.nepjol.info/index.php/rnjds/article/view/22425/19081.

[4] Tek Bahadur Thapa, Agricultural Extension Services Delivery System in Nepal, Food and Agriculture Organization of the United Nations, UN Complex, Pulchowk, Nepal, June 2010, https://reliefweb.int/sites/reliefweb.int/files resources/AgExtServDelSysNepal.pdf.

[5] Bashu Dev Deshar, An Overview of Agricultural Degradation in Nepal and its Impact on Economy and Environment, Global Journal of Economic and Social Development. Volume 3, Number 1 (2013) pp. 1-20 © Research India Publications, https://www.researchgate.net/publication/320146140_An_Overview_of_Agricultural_Degradation_in_Nepal_and_its_Impact_on_Economy_and_Environment

[6] FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS, Challenges and opportunities for Nepal's small andmedium forest enterprises(SMFEs)ForestConnectForest Connect Diagnostic Studies on Small and Medium-Sized Forest Enterprises, Asia Network for Sustainable Agriculture and Bio resources Kathmandu, 2009, Nepal, http://www.fao.org/3/i1266e/i1266e00.pdf.

[7] David Kahan, Market-oriented advisory servicesin Asia – a review and lessons learned, RAP PUBLICATION 2011/10, FOOD AND AGRICULTURE

ORGANIZATION OF THE UNITED NATIONSREGIONAL OFFICE FOR ASIA AND THE PACIFIC, Bangkok, 2011, http://www.fao.org/3/a-i2324e.pdf.

[8] Ministry of Frogien Affairs, Nepal: Market Study Identifying Foreign Direct Investment Opportunities in Nepal, Commissioned by the Netherlands Enterprise Agency, 1 May 2018, https://www.rvo.nl/sites/default/files/2018/10/Nepal-Market-Study.pdf

[9] Deepak M. Pokhrel, Gopal B. Thapa, Are marketing intermediaries exploiting mountain farmers in Nepal? A study based on market price, marketing margin and income distribution analyses, Agricultural Systems,Volume 94, Issue 2,ISSN 0308-521X, doi.org/10.1016/j.agsy.2006.08.004.

[10] Mina Nath Paudel,Prospects and limitations of agriculture industrialization in Nepal, Agronomy Journal of Nepal (Agron JN), Vol.4, 2016, www.nepjol.info/index.php/AJN/article/view/15515/12519.

[11] T.L. Thapa et al. Supply Chain Analysis of Rice Sub-Sector in Dang District, Nepal, (2018) Int. J. Appl. Sci. Biotechnol. Vol 6(4): 319-326, DOI: 10.3126/ijasbt.v6i4.22109

[12] FAO Representation, Nepal and FAO Achievements and success stories, May 2011, Rome, Italy, http://www.fao.org/3/a-at013e.pdf.

[13] OECD, Adoption of Technologies for Sustainable Farming Systems Wageningen Workshop Proceedings, OECD Publications, 2001, 2, rue André-Pascal, 75775 Paris Cedex 16, France, https://www.oecd.org/greengrowth/sustainable-agriculture/2739771.pdf

[14] FROST & SULLIVAN and Government of Nepal, Ministry of Communication and Information Technology, 2018 Digital Framework Nepal, © 2018 FROST & SULLIVAN, https://mocit.gov.np /application/resources/admin/uploads/source/EConsultation/Final%20Book.pdf

[15] B. Devkota, B. Adhikar and D. Shrestha, "Integrating romanized Nepali spellchecker with SMS based decision support system for Nepalese farmers," 2015 9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), Kathmandu, 2015, pp. 1-6, doi: 10.1109/SKIMA.2015.7400046.

[16] Craig Larman, Applying UML and Patterns, Pearson Education, 3rd Edition, 2005, ISBN-13: 978-0131489066, ISBN-10: 0131489062

[17] Roger S Pressmen, Bruce Maxim, Software Engineering: A Practitioners Approach. Tata McGraw Hills, 8th Edition, ISBN-13: 978-93-5316-571-0, LLC, New York

Automation Mechanism of Detecting Suspected Piracy Video Streaming Sites

Minkyung Lee[1], Deuk-Hun Kim[1] and Jin Kwak[2*]
[1]ISAA Lab., Department of Computer Engineering, Ajou University
Suwon, Republic of Korea
[2]Department of Cyber Security, Ajou University
Suwon, Republic of Korea
[e-mail: leemkl@ajou.ac.kr, kimdh1206@ajou.ac.kr, security@ajou.ac.kr]
*Corresponding author: Jin Kwak

Abstract

Recently, watching videos such as movies and dramas is increasing by using video streaming services rather than video downloading services worldwide. As a result, the scale of a video streaming market is increasing. However, video streaming sites that infringe copyrights such as movies and dramas are also on the rise. In response to this, it takes action such as closing piracy streaming sites, but the sites are becoming more sophisticated and intelligent to bypass it. Even intelligent piracy sites are almost identical to legal streaming sites. So, it could not be detected by existing piracy streaming site feature-based detection techniques. Therefore, in this paper, we analyze features of legal streaming sites and suspected piracy video streaming sites that cause copyright infringement. Also, we propose a technique of detecting suspected copyright infringement video streaming sites based on a whitelist to respond to increasing and intelligent piracy sites.

***Keywords*:** Copyright, Copyright infringement, Piracy sites, Whitelist

1. Introduction

Digital contents such as films and TV programs can be accessed more easily through web browsers as IT(Information Technology) advances. Recently, however, there have been piracy sites such as unauthorized distribution or reproduction of digital contents without the the copyright holder's permission. Besides, they distribute digital content in various types, such as downloading, torrent, and streaming [1]. Among them, video streaming service is increasing rather than directly downloading videos, and piracy video streaming sites are also increasing [2]. There are also increasing advanced piracy video streaming sites similar to legal video streaming sites and not characteristic of general piracy video streaming sites. Therefore, studies are needed to detect automatically in order to efficiently detect piracy video streaming sites and are to detect advanced piracy video streaming sites that are similar to legal video streaming site. Accordingly, in Chapter 2, we analyze features of piracy video streaming sites and detecting methods of piracy sites in the existing studies. In Chapter 3, we propose an automation mechanism of detecting suspected piracy video streaming sites using a whitelist to detect advanced piracy video streaming sites. Finally, we present the conclusion in Chapter 4.

This research project was supported by Ministry of Culture, Sports and Tourism(MCST) and from Korea Copyright Commission in 2020(2019-PF-9500).

2. Related Works

2.1 Feature of piracy video streaming sites

Feature of piracy video streaming sites for detecting piracy video streaming sites that infringe on digital content are as follows [1].

• Content list

The piracy video streaming sites include a content list that contains the title of content so that users can easily access to watch digital content.

• Advertisement

The piracy video streaming sites include advertising banners in various parts for profit-making reasons.

• Streaming server list

In some cases, the piracy video streaming sites also include another streaming server list without embedded streaming video content. Even if one streaming server is blocked, a list of two or more streaming servers can be provided for the user to select the desired quality.

2.2 Methods of detecting piracy sites

The proposed methods for detecting piracy sites in previous studies are as follows.

❒ Feature analysis and detection technique for piracy sites

The piracy site was classified as a torrent, video streaming and webtoon site, and feature analysis was performed in this technique. Subsequently, the torrent, video streaming, and webtoon sites were detected as piracy sites according to each piracy site [1].

❒ A detection of webtoon infringement site using page source analysis

This technique can detect webtoon infringement sites through page source analysis by crawling websites. This technique derived a business registration number in legal sites and keyword characteristics used in webtoon infringement sites. As a result, they can detect suspected infringement sites according to the keyword list. In comparing and analyzing the keyword list, the accuracy was increased using the process of separating compound words rather than simply comparing them [3].

❒ Detection technique of suspected webtoon infringement site using web site advertisement banners

This technique can detect webtoon infringement sites by analyzing advertising banners in piracy websites, a profit-making element of piracy sites. After extracting websites in image form through Full Page Screen Capture, a google extension program, they used gambling-related keyword lists used in piracy sites through OCR(Optical Character Recognition) analysis to distinguish legal sites from piracy sites [4]. Piracy video streaming sites also utilize advertising banners in web sites, which are a profit-making factor, and the advertising banners are the same as the keyword list used by the piracy sites. Therefore, we propose an automation mechanism using the analysis of advertising banners.

2.3 Whitelist

We use a whitelist technique in the proposed mechanism for detecting piracy video streaming sites in this paper. The whitelist is used as a positive security method compared the blacklist, which allows only proven ‘security’ and blocks proven ‘malicious’ [5]. As piracy video streaming sites are advanced and intelligent with similar features to legal streaming sites, advanced piracy video streaming sites have difficulty detecting piracy sites using features of legal video streaming sites and piracy video streaming sites. Therefore, we apply a whitelist technique to detect or monitor malware, phishing sites, etc. to detect advanced piracy video streaming sites.

3. Proposed Scheme

3.1 Analysis of legal and piracy video streaming sites

We organize the legal video streaming site and piracy video streaming site features as shown below in **Table 1**. Features of sits used in **Table 1** are as follows:

• Business Registration Number

The legal video streaming site includes the business registration number, which is the National Tax Service’s management number.

• Contents list

Name of contents for providing digital contents on the video streaming sites

• Banner Advertising

Gambling-related advertising banner

• Streaming Server list

A link list to another external streaming site

• Embedded Video

Video to watch digital content directly from a site without moving to an external site

Table 1. Analysis of comparison legal video streaming sites to piracy video streaming sites

Feature of sites	Legal sites	Piracy sites	Advanced Piracy sites
Business Registration Number	O	X	O (Invalid number)
Content lists	O	O	O
Banner Advertising	X	O	X
Streaming Server list	X	O	X
Embedded video	O	X	O

While piracy video streaming sites contain streaming server lists and advertising banners, there are advanced piracy sites that do not include streaming server lists and advertising banner. They also include a business registration number and an embedded video similar to the feature of legal video streaming sites, such as **Table 1**.

3.2 Feature of piracy video streaming sites

❒ Streaming keyword

Table 2. List of streaming keyword using piracy sites

Video Keyword	Piracy Site		
	ogu*** .com	the*** .guru	joo*** .net
Replay link	O	O	-
Only link	O	O	O
Legal issue	O	O	-
Copyright	O	O	O
Provided link	O	O	-
Inquiry of advertising	-	O	-

Streaming keywords, which are mainly used in video streaming sites, are shown in **Table 2**.

❒ Link list keyword

Table 3. List of link keyword using piracy sites

Link List Keyword	Piracy Site		
	ogu*** .com	the*** .guru	joo*** .net
Glimited	-	O	O
HPlay	-	O	O
JawCloud	-	O	-
Jload	-	O	O
MixDrop	O	-	-
JetLoad	O	-	-
FlashVid	O	O	-
SuperVid	O	O	O

Piracy site provides streaming services by providing external streaming sites. Therefore, a link list used in piracy sites is derived to extract commonly used external streaming sites. The link list, which is commonly used on the three piracy video streaming sites, is shown in **Table 3**.

3.3 Proposed automation mechanism for detecting piracy video streaming sites

In this paper, we propose a whitelist-based automation mechanism of detecting piracy video streaming sites to detect advanced piracy video streaming sites. Accordingly, the analysis of legal and piracy video streaming sites and the feature of piracy video streaming sites are utilized, and the mechanism proposed in this paper are as follows, **Fig. 1**.

Step 1. Site Crawling and Suspicion

Kim et al. proposed the method of site crawling and checking a business registration number. Unlike the existing proposed method, piracy video streaming sites are advanced and intelligent, similar to legal video streaming sites. Therefore, it is difficult to determine that it is not a piracy video streaming site based on whether business registration numbers are included. If the business registration number is included in the site, it goes to *Step 4*. The pseudo-code for implementing site crawling and checking a business registration number is shown in **Table 4**.

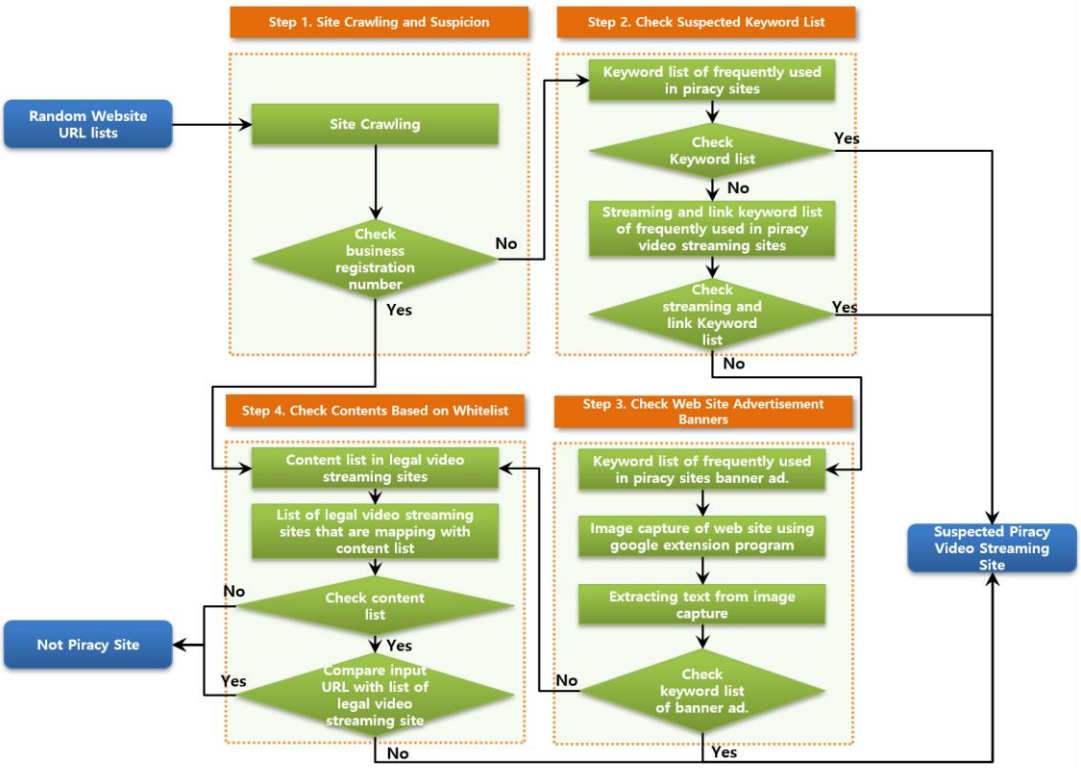


Fig. 1. Proposed mechanism for detecting piracy video streaming sites

Table 4. Pseudo-code for site crawling and checking suspicion

Function for site crawling and checking suspicion:	
1	*str_html* ← crawled site
2	*license_signatrue* ← regular expression for determining business registration number
3	IF there is business registration which is matching with *license_signature* in a site,
4	RETURN business registration number
5	ELSE, RETURN -1

Step 2. Check Suspected Keyword List

This step compares to the keyword list mainly used in the piracy sites used in the proposed method by Kim et al. The following Table 5 is a pseudo-code for *Step 2*.

Table 5. Pseudo-code for checking keyword list

Function for checking keyword list:	
1	*str_html* ← crawled site in *Step 1*
2	*commonKeyword* ← keyword list mainly used in piracy site
3	FOR keyword in *commonKeyword*:
4	IF keyword is mapped by *str_html*, RETURN 0
5	RETURN -1

Besides, the keyword of the piracy video streaming sites derived from this paper is applied in *Step 2*. This step compares with the streaming keyword list(streaming keyword in Table 2) and the link keyword list(link list keyword in Table 3) based on features of piracy video streaming sites. The following Table 6 is a pseudo-code for implementing *Step 2*.

Table 6. Pseudo-code for checking streaming and link keyword list

Function for checking streaming and link keyword list:	
1	*str_html* ← crawled site in *Step 1*
2	*streamingKeyword* ← streaming keyword list
3	*linkKeyword* ← link keyword list
4	FOR keyword in *streamingKeyword*:
5	IF keyword is mapped by *str_html*, RETURN 0
6	FOR keyword in *linkKeyword*:
7	IF keyword is mapped by *str_html*, RETURN 0
8	RETURN -1

Step 3. Check Web Site Advertisement Banners

It can check advertisement banner in piracy video streaming site using the proposed method by Kim et al. The following **Table 7** is a pseudo-code for implementing *Step 3*.

Table 7. Pseudo-code for checking banner advertising

Function for checking banner advertising:	
1	*Ad_keyword* ← keyword list of banner advertising
2	*img* ← image capture using a google extension program
3	*img_text* ← converting *img* to text
4	FOR keyword in *Ad_keyword*:
5	IF keyword is mapped by *str_html*, RETURN 0
6	RETURN -1

Step 4. Check Contents Based on Whitelist

It proposes a mechanism based on a whitelist for sites(false negative) that have not been detected as suspected piracy video streaming sites through the results of *Step 1* to *Step 3*. It prepares a list of the content list that was broadcasted and posted on the legal video streaming sites before using a whitelist. It also prepares a URL(Uniform Resource Locator) list of legal video streaming sites, which are mapping with a prepared list of the content. After that, it compares a content list in not suspected piracy video streaming sites to the prepared content list. If there is no content in a site, it detects not suspected video streaming sites. If there is content in a site and the URL of the site is not legal video streaming sites, it detects suspected video streaming site. The following **Table 8** is a pseudo-code for implementing *Step 4*.

Table 8. pseudo-code for detection piracy sites using a whitelist

Function for detection piracy sites using whitelist:	
1	*str_html* ← crawled site in *Step 1*
2	*contents* ← content list that is posted in legal video streaming sites
3	*legal_list* ← legal video streaming site URL corresponding to *contents*
4	While(repeat by the length of *contents*):
5	IF there is the same content as the *contents* in str_html,
6	IF URL of str_html does not match in the *legal_list*, RETURN -1
7	ELSE RETURN 0
8	RETURN -1

In this paper, we proposed the mechanism, which applied features of piracy video streaming site and the technique of previous studies to detecting piracy sites in *Step 1* to *Step 3*. In addition, we applied whitelist to sites that were not detected as sites suspected of piracy video streaming sites in *Step 1* to *Step 3* to improve the detection rate of suspected piracy video streaming sites.

4. Conclusion

Various piracy sites can detect suspected piracy sites using common features because piracy sites are similar to each other. Besides, the detection rate can be improved by applying a whitelist technique if the feature of the piracy sites is similar to those of legal sites due to the sophistication and intelligence of piracy sites. In the future, we will study ways to improve the detection rate of suspected piracy sites of various types of infringement such as torrents and webtoons through whitelist-based techniques and apply high-speed search algorithms to improve the detection performance of suspected piracy sites.

References

[1] S.K. Choi and J. Kwak, “Feature Analysis and Detection Technique for Piracy Sites”, *KSII Transactions on Internet and Information Systems*, Vol 14, No. 5, pp 2204-2220, May. 2020.

[2] European Union Intellectual Property Office (EUIPO), “Online Copyright Infringement in The European Union: Music, Film, TV (2017-2018), Trends and Drivers”, *European Union Intellectual Property Office*, Nov. 2019.

[3] E.J. Kim, S.K. Choi and Jin Kwak, “Detection Technique of Suspected Webtoon Infringement Site using Web Site Advertisement Banners”, *Proceedings of KSII Spring Conference*, May. 2020.

[4] E.J. Kim, S.K. Choi and J. Kwak, “A Detection of Webtoon Infringement Site Using Page Source Analysis”, *Proceedings of the 15st Asia Pacific International Conference on Internet Science and Technology (APIC-IST 2020)*, Jul. 2020.

[5] A. Sedgewick, M. Souppaya and K. Scarfone, “Guild to Application Whitelisting”, *National Institute of Standards and Technology (NIST)*, Oct. 2015.

Prediction Method for New Domain of Suspected Copyright Infringement Site

Hae-Seon Jeong[1], Deuk-Hun Kim[2] and Jin Kwak[3*]
[1]ISAA Lab., Department of AI Convergence Network, Ajou University
Suwon, Republic of Korea
[2]ISAA Lab., Department of Computer Engineering, Ajou University
Suwon, Republic of Korea
[3]Department of Cyber Security and Department of AI Convergence Network, Ajou University
Suwon, Republic of Korea
[e-mail: haeseon0363@ajou.ac.kr, kimdh1206@ajou.ac.kr, security@ajou.ac.kr]
*Corresponding author: Jin Kwak

Abstract

A suspected copyright infringement site, which is suspected of illegally sharing copyrights and violating copyright holders' rights, is monitored by the Korea Copyright Protection Agency to reduce copyright holders' damage and protect copyright. However, suspected copyright infringement site bypass monitoring by using new domains to change their access address. Therefore, efficient monitoring of suspected copyright infringement sites requires predictions of new domains to be created and the operation start date of the site using a new domain. However, the frequency from creating new domains to operating sites varies from site to site. Also, it is difficult to identify the duration of the closed site, which has a previous domain, because it is no longer accessible. However, you can determine the start date of the operation by using an index search which prints the posting date of sub-webpages. Also, as suspected copyright infringement site follows the rule of creating a domain, the new domain can be predicted. Therefore, in this paper, we increase the efficiency of monitoring by proposing a method of predicting new domain and the start date of operation by utilizing the results of the index search.

***Keywords*:** Copyright, Suspected copyright infringement site, Prediction method

1. Introduction

Copyright infringement sites illegally share works without permission from copyright holders. As the demand for copyright infringement sites continues to grow, with free access to copyrighted works, the extent of copyright holders' damage is also increasing.

A site that is suspected of illegally sharing works and needs to be monitored is called a suspected copyright infringement site. To block the operation of suspected copyright infringement sites and the use of a site by users, the Korea Copyright Protection Agency lists the domains of suspected copyright infringement sites and monitors those domains continuously [1].

Copyright infringement site has steps over time which is defined as lifecycle to respond to monitoring. Suspected copyright infringement site responds to monitoring by blocking access requests or performing redirection to another site. Also, it responds to monitoring by changing the access address to a new domain. However, even if a suspected copyright infringement site block access requests, monitoring can be carried out by bypassing them using VPN, etc. Also, if a suspected copyright infringement site is

This research project was supported by Ministry of Culture, Sports and Tourism(MCST) and from Korea Copyright Commission in 2020(2019-PF-9500).

redirected to another site after a request for access, monitoring can be performed by analyzing the response domain.

However, if a suspected copyright infringement site creates a new domain and uses it to operate the site, it will be difficult to secure a target domain to monitor. Monitoring can be carried out quickly if the new domain of the suspected copyright infringement site and the operation start date of the site using it can be predicted. Each site has a difference in the cycle from the creation of a new domain to the start date of the site's operation, which can be predicted from the start date of previous domain operations. However, it is difficult to obtain information about the operation start date of the previously closed domain using a search. Although, you can obtain a variety of information within sub-webpages of the site which is already closed by using index search. Thus, it is possible to predict a new domain based on the domain creation rules of copyright infringement sites and to predict the operation start date based on index search results.

Accordingly, in this paper, the monitoring response techniques are analyzed through the lifecycle of the suspected copyright infringement site, and the domain creation rules and index search results of the suspected copyright infringement site are analyzed in Chapter 2. In Chapter 3, we propose a method to predict new domain, creation date, and operation start date using domain history and finally concludes in Chapter 4.

2. Related Works

2.1 Analysis of Lifecycle Model

The lifecycle model for suspected copyright infringement sites is defined as five steps of creation, operation, response, change, and closure as shown in **Table 1** [2]. The creation is the step where the domain is first registered and the site is created, and the operation is the step where the site is operated providing copyright infringement contents. The action is a step that allows only bypass access to both the latest and previous versions of the domain. The change is a step when it is redirected to another suspected infringement site. Action and change are steps to respond to monitoring, but monitoring can be performed if you access through VPN or response domains are identified. The closure is a step where services are no longer provided and operations are discontinued.

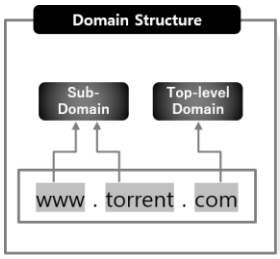
Table 1. Steps of lifecycle model

Step	Description
Creation	The Site is created through domain registration
Operation	Operators or users provide copyright infringement contents
Action	Bypass access is available when the domain of a suspected copyright infringement site is updated to the latest version.
	If the domain of a suspected copyright infringement site is not updated to the latest version, normal access is not possible but only bypass access is possible.
Change	When the domain of the suspected copyright infringement site is updated to the latest version, it is accessible without bypassing it and then redirected to another site
	When the domain of the suspected copyright infringement site is not updated to the latest version, normal access is not possible, but only bypass access is available and redirected to another site
Closure	The site does not provide copyright infringement contents anymore and no more operates the site

However, to prepare site blockage and monitoring, the suspected copyright infringement site creates a new domain and use the domain to resume operation of the site. Therefore, monitoring suspected copyright infringement site requires predicting new domains and predicting the start date of operation of the sites that use a new domain.

2.2 Analysis of Rules for Generating Domain

Many suspected copyright infringement sites generate new domains by following the rules in **Fig. 1**. A domain consists of a sub-domain and a top-level domain, depending on the location of the elements created when separate a domain by ".". The element on the right becomes the top-level domain, and the element on the left becomes the sub-domain. There may be several sub-domains, and the more they are located on the left, the lower the sub-domain. Suspected



# of Sub-Domain	Rules for Generating New Domain	Example	
		Current Domain	New Domain
1	Add new sub-domain	torrent.com	www.torrent.com
	Change top-level domain	torrent.com	torrent.site
	Change/Remove number of sub-domain	torrent.com	torrent6.com
2	Change/Remove new sub-domain	www.torrent.com	torrent.com
	Change top-level domain	www.torrent.com	www.torrent.me
	Change/Remove number of sub-domain	www.torrent.me	www.torrent2.me

Fig. 1. Rules for generating domain

copyright infringement site generates a new domain by changing the top-level domain or deleting sub-domain. Also, if a number comes before or after a sub-domain, increase the number to create a new domain. You can create a new domain by mixing two or more of the new domain creation rules [3]. Therefore, domain generating rules allow you to predict past domains or predict new domain.

2.3 Analysis of Index Search

The copyright infringement site has a different operating cycle for each site. To know the operating cycle of a site, you need information about the start date of the site's operation in the previous domain. However, for previous domains that are currently closed and inaccessible, it is difficult to retrieve relevant information on the start date of operation. However, Google's index search function allows users to collect information on closed sites about the start date of the site's operation in previous domains. Google performs indexing on sub-webpages of multiple domains. Google provides an index search with the "site:" operator which prints information about the sub-webpage indexed [4].

Suspected copyright infringement site posts the name of the post and the date of posting on the webpage when uploading the work as shown in **Fig. 2**. The following is a detailed description of the information that exists within the result of the index search.

• ":site" Operator

Perform index search about a domain

• Search Result

Saved information when the domain was indexed

• Sub-webpage

sub-webpage with URL starting with the searched domain

Fig. 2. Result of index search

• Posting Date

The Posting date of a work of art in sub-webpage with URL starting with the searched domain

The posting date means that the domain searched was operated during the period, so the earliest posting date can confirm the operation start date of the site using the domain.

3. Proposed Method

This chapter proposes a method to predict the new domain and the operation start date of the site, based on the domain history of the suspected copyright infringement site and the cycle of the start date of operation utilizing it.

In section 3.1, We generate a list of domain history for suspected copyright infringement sites based on domain generation rules. In section 3.2, we analyze the domain creation date and the site's operation start date for the domain change list. Based on this, in section 3.3, we propose a method to predict the new domain, the creation date of it, and the start date of site operation using it. Details of the proposed method are as follows.

3.1 Generation of Previous Domain History

Create a domain history first to analyze the creation dates and start dates of the previous domains of suspected copyright infringement site.

By using the domain creation rules analyzed in Section 2.2, we generate a list of possible domain changes. The detailed process for creating a domain history of a suspected copyright infringement sites is as follows.

Step 1. Decrease/Delete Number in Sub-domain

The suspected copyright infringement site creates a new domain by increasing the number in the sub-domain. Therefore, it can be expected that a number smaller than the current number in the sub-domain exists within the previous domain, or that a number does not exist. Therefore, the domains created by reducing the number in the sub-domain by 1 are added to the domain history and the domain that has removed the number in the sub-domain also added to the domain history as shown in **Fig. 3**.

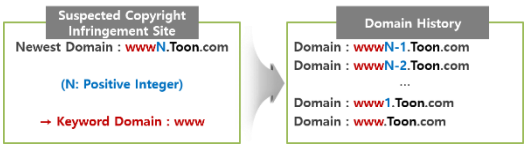


Fig. 3. Example of domain history when two sub-domains are present in the current domain

Step 2. Remove Sub-domain of Left-side

The suspected copyright infringement site creates a new domain when it has one sub-domain by adding a new sub-domain on the left-side. Therefore, if two sub-domains exist, it can be expected that no sub-domain on the left would have existed in the previous domain. Therefore, if there are two sub-domains, the domain that has removed the sub-domain located on the left-side is added to the domain history as shown in Fig. 4.

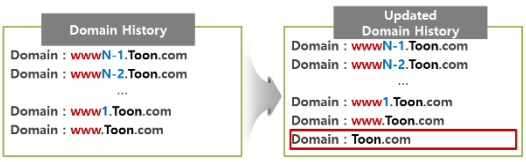


Fig. 4. Example of Domain history updating

Step 3. Check Top-level Domain

The suspected copyright infringement site creates a new domain by changing the top-level domain. Therefore, after combining the top-level domains used by suspected copyright infringement sites, only the domain where the index search results exist is added to the domain history as shown in **Fig. 5**.

After generating a domain change history through the above process, we analyze the domain creation date and site operation start date.

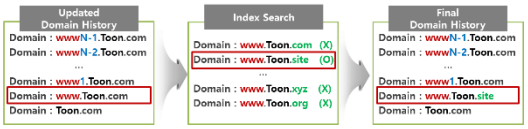


Fig. 5. Process of top-level domain check

3.2 Analysis of Domain Operation Period

Each suspected copyright infringement site has a different cycle from creating a new domain to using it for operation. Therefore, to predict the operation start date of a site using a new domain, an analysis of the domain creation date and the start date of operation of the domains in the history is required. The detailed analysis process for this is as follows.

Step 1. Check Domain Creation Date

As the domain creation date can be checked in the "Creation Date" of the WHOIS response, we request the domain creation date from the WHOIS server. The example of **Fig. 6** shows that the creation date for a domain with increased number is up to date.

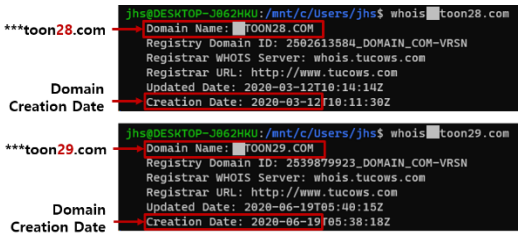


Fig. 6. Domain creation date in whois response

Step 2. Check Posting Date by Index Search

By performing an index search with the ":site" operator, you can check the posting date of posts within the suspected copyright infringement site. Therefore, the posting date can be determined by matching the regular expression which means the date for the index search results. The earliest date of the results is identified as the start date of operation.

3.3 Prediction Method for Next Domain

Based on the creation date and the operation start date of the previous domains, we make predictions about the new domain that will be created in the future, as well as the creation date and start date of operation. We secure the target list to monitor first and then monitor a new domain during the predicted period.

When the creation of a new domain is detected, we monitor the operation start of the suspected copyright infringement site during the predicted period. The symbols for predicting the creation date of the domain and the operation start date of the site are shown in **Table 2**. The detailed process of predicting the method is as follows.

Table 2. Symbols for prediction of new domain use

Symbol	Description
$N-1$	The number of previous domains
i	The sequence number of a domain $(1 \le i \le N-1)$
c_i	Creation date of a previous domain $(1 \le i \le N-1)$
o_i	Operation date of a previous domain $(1 \le i \le N-1)$
c_N	Creation date of a current domain
o_N	Operation start date of a current domain
c_{N+1}	Creation date of a new domain
o_{N+1}	Operation start date of a new domain
cd_i	Difference between the domain creation dates, $c_{i+1} - c_i$
od_i	Difference between the domain creation date and the operation start date, $c_i - o_i$
m_c	Mean of cd_i, $\frac{\sum_{i=1}^{i=N-1} cdi}{N-1}$
m_o	Mean of od_i, $\frac{\sum_{i=1}^{i=N} odi}{N}$
σ_c	Standard deviation of cd_i, $\sqrt{\frac{\sum_{i=1}^{i=N-1}(cd_i - m_c)}{N-1}}$
σ_o	Standard deviation of od_i, $\sqrt{\frac{\sum_{i=1}^{i=N}(od_i - m_o)}{N}}$
P_c	The predicted range of c_{N+1} $(c_N \le P_c \le c_N + m_c + \sigma_c)$
P_o	The predicted range of o_{N+1} $(c_{N+1} \le P_o \le c_{N+1} + m_o + \sigma_o)$

Step 1. Predict New Domain List

The process of creating a new domain list is similar to the process of creating a domain history in Section 3.1. Generate a list of new domains, by increasing the current number A $(1 \le A)$ in the sub-domain numbers from $A+1$ to $A+10$. If there are two sub-domains, add a new domain that removed the left-side sub-domain to the new domain list. Next, for each new domain, the domains created by changing the top-level domains used by suspected copyright infringement sites are added to the new domain list.

Step 2. Predict Domain Creation Date

For the history domain changes of one suspected copyright infringement site, obtain the mean (m_c) of the differences (cd_i) between the domain creation dates$(c_i)$and obtain the standard deviation (σ_c), which is the average for the difference between cd_i and m_c. The new domain creation date (c_{N+1}) may be the same as the current version of the domain creation date (c_N), and there is a high probability that a new domain will be created for a period up to $c_N + m_c + \sigma_c$ based on the domain history. Therefore, the period from c_N to $c_N + m_c + \sigma_c$ is set to the predicted range of new domain creation date(P_c). We monitor the creation of a new domain through WHOIS for the period corresponding to P_c. If the new domain creation is detected, monitor the start of site operations using the new domain through the following process.

Step 3. Predict the Operation Start Date

If creation of the new domain is detected, the site operation is monitored for a period of that range by predicting the scope of the start date of the site's operation.

For the history domain changes of one suspected copyright infringement site, obtain the mean (m_o) of the differences(od_i) between the domain creation date(c_i) and the operation start date(o_i) and obtain the standard deviation(σ_o), which is the average for the difference between od_i and m_o. The operation start date of the new domain (o_{N+1}) may be the same as the creation date of the new domain (c_{N+1}), and there is a high probability that the operation start date of the new domain will be created for a period up to $o_N + m_o + \sigma_o$ based on the domain history. Therefore, the period from c_{N+1} to $c_{N+1} + m_o +$

σ_o is set to the predicted range of operation start date of the new domain (P_o). We monitor and detect the operation start through characteristics of the suspected copyright infringement sites.

4. Conclusion

The suspected copyright infringement site responds to monitoring by generating a new domain and using it as an access address for the site. Thus, to monitor suspected copyright infringement sites efficiently, new domain and operation start date using it must be predicted.

In this paper, we predicted a new domain name through the rules for generating domain and predicted the new domain creation date and the operation start date by using result of index search.

By using this prediction method, we can quickly detect the start of a suspected copyright infringement site operation using the new domain. Thereby, we can minimize the unauthorized work-sharing and the resulting damage to copyright holders.

In the future, we suppose to improve the accuracy of previous domain history by subdividing the rules for adding sub-domain when forming a new domain. Also, we plan to experiment on the proposed prediction method to verify the predicted results.

References

[1] Korea Copyright Protection Agency, “2020 Annual Report on Copyright Protection”, *Korea Copyright Protection Agency*, pp. 1-174, Apr. 2020.

[2] D.H. Kim, H.S. Jeong and J. Kwak, “Development of Lifecycle Model for Copyright Infringement Site”, *Journal of The Korea Institute of Information Security & Cryptology*, Vol. 30, No. 1, pp. 101-121, Feb. 2020.

[3] H.S. Jeong and J. Kwak, “Automatic Detection Method for Life Cycle Step of Suspected Copyright Infringement Site”, *Proceedings of Summer Conference on Information Security and Cryptography*, Vol. 30, No. 1, pp. 477-480, Jul. 2020.

[4] Google Search Quality team, “Search Engine Optimization (SEO) Starter Guide”, *Google*, 2020.

Automated Detection Architecture for Piracy Sites Based on Link Collection Sites

Eui-Jin Kim[1] Deuk-Hun Kim[1] and Jin Kwak[2*]
[1]ISAA Lab., Department of Computer Engineering, Ajou University
Suwon, Republic of Korea
[2]Department of Cyber Security, Ajou University
Suwon, Republic of Korea
[e-mail: dmlwls0403@ajou.ac.kr, kimdh1206@ajou.ac.kr, security@ajou.ac.kr]
*Corresponding author: Jin Kwak

Abstract

With the recent development of smart devices and media services, people can use digital content, including movies, dramas, and music. The scale of the domestic copyright market continues to increase, and the scale of overseas copyright exports is also increasing due to the overseas export of K-Contents. However, as the copyright market scale has increased, piracy sites have emerged, and piracy sites have also increased. To prevent this, public organizations are in the process of determining whether there is a suspicion of copyright infringement and blocking piracy sites, but there are difficulties in finding the continuously created piracy sites. Moreover, the new piracy sites and existing piracy sites are updated on the link collection sites and provided to clients. Therefore, in this paper, I propose the advanced automatic detection architecture of piracy sites based on the link collection site.

***Keywords*:** Copyright, Piracy Site, Link Collection Site

1. Introduction

Recently, media services have been developing due to of IoT(Internet of Things) technologies, including smart devices, and the 'snack culture' that consumes cultural content in a short time [1]. As a result, individuals have increased their access to digital content, and the copyright market is also increasing. Also, due to the overseas expansion of webtoon-based K-drama and K-films, the scale of the copyright market is increasing due to the export of webtoons, movies, and dramas to overseas markets [2]. With the increase in the scale of the copyright market, piracy site's scale has increased, and the damage caused by piracy sites has also increased [3]. To prevent damage caused by copyright infringement, public institutions are making efforts to block copyright infringement sites. Still, it is difficult to prevent copyright infringement because the infringement site operator continues to operate the site through domain changes or creates and operates a new infringement site [4]. These domains changed piracy sites, and newly created piracy sites are updated on link collection sites and provided to clients, but research on link collection sites is insufficient.

In this paper, chapter 2 deals with the existing research on site detection techniques, and chapter 3 deals with the feature of link collection sites. Chapter 4 proposes advanced automated detection architecture for piracy sites, chapter 5 deals with the analysis of the proposed techniques, chapter 6 deals with conclusions.

This research project was supported by Ministry of Culture, Sports and Tourism(MCST) and from Korea Copyright Commission in 2020(2019-PF-9500).

2. Site Detection Techniques

2.1 Detection Technique for Blacklist Site

Blacklist detection techniques warn Web browser users when they access blacklisted websites. Blacklist detection techniques have the advantage of being easy to implement algorithms and low detection rates. Solutions using the Blacklist Detection Method include Google's ‘Safe Browsing’ and Microsoft's ‘Phishing Filter’. Blacklist detection techniques were tested on phishing sites by researchers at the University of Vienna, showing a high detection rate of 90.23% [5]. However, there is a disadvantage of not detecting sites that are not registered, and they are taking follow-up measures such as blacklisting them after blocking the site [6]. It is also difficult to detect blacklists in Korea due to the lack of its blacklists collection and sharing system compared to overseas countries [5].

2.2 Search Technique for Similar Domain Site

A similar domain search technique is a detection technique that uses registering a similar domain when creating a piracy site using domain address keywords and is used when there is a certain pattern in the domain. However, a disadvantage cannot be detected if the Zone file of the Top-Level Domain is not open to the outside world or if the domain pattern changes. Johns Hopkins University and Google researchers tested 2,508 URL(Uniform Resource Locator)s with 'Local Regression Technique', 'Domain-based Feature Analysis', and 'URL Keyword Comparison Method'. As a result, 95.8% of the detection rate and 1.2% of the false detection rate were demonstrated to be effective. However, even a slight change in the URL pattern can reduce the detection rate [5].

Existing detection methods do not reflect new infringement sites and lower detection rates when changing the infringing site’s domain. On the other hand, the link collection site already reflects the new and existing infringement site’s domain changes. Therefore, it is necessary to research link collection sites.

3. Feature of Link Collection Site

In this chapter, we describe the features of the link collection site. **Fig. 1** shows an example of a link collection site.

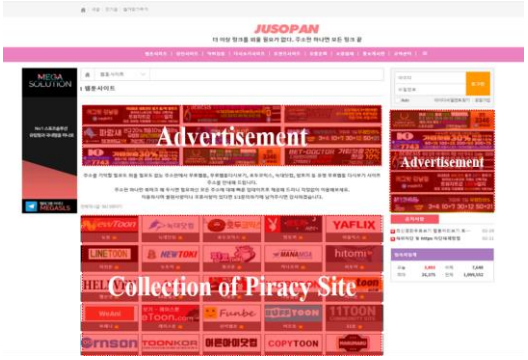


Fig. 1. Example of link collection site

Link collection sites are sites that list and provide piracy sites to clients. Link collection sites generate revenue by posting illegal advertising banners inside the site, and some piracy sites exist in the form of banners on the website's internal frame. Accessing the URL associated with the piracy site banner links to a page that can derive the piracy site’s name and domain address called the Internal URL. **Fig. 2** represents the Internal URL.

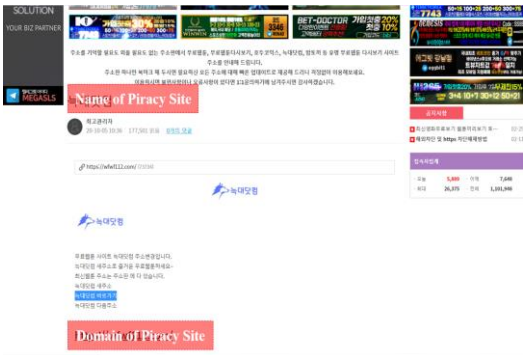


Fig. 2. Example of internal URL

Link collection sites offer Torrent, video streaming, and webtoons as well as pornography and gambling, and they post site links reflecting domains that change periodically due to the short life cycle of the piracy site. Also, new piracy sites and previously closed sites, are being updated on link collection sites.

Fig. 3. Proposed scheme overview

4. Proposed Scheme

This chapter, proposes advanced automated detection architecture for piracy sites using link collection sites, and **Fig. 3** is a picture of the proposed scheme.

Step 1. Access Link Collection Site

First, access the link collection site. Using the features of multiple infringement sites in the form of banners on the link collection site page, find the 'a href' tag that connects to the banner link inside the frame. Exclude the URL corresponding to the tag 'a *href*' from links to the external URL other than the link collection site. The external URL includes an ad banner link, a notice board link, etc., and extracts an internal URL that can lead to piracy sites by conducting a refinement process that excludes external URLs.

Step 2. Access Internal URL of Link Collection Site

Access the internal URL derived from *Step 1* and derive the URL accessible directly to the infringing site. This process uses the features in which the piracy site's name and domain address exist on the Internal URL page.

Step 3. Create Data Set of Piracy Site

The data set is created by storing the piracy site's domain address derived from *Step 2* in the 'csv' format. This process derives all the data sets for torrent, video streaming, and webtoon piracy sites, which database piracy sites.

5. Analysis of Proposed Scheme

In this chapter, we deal with the experimental results of the proposed scheme in chapter 4. The experimental environment for implementing the proposed scheme is, as shown in **Table 1**. **Fig. 4** is a code that implements the proposed scheme.

Table 1. Experiments environment

Operating System	Windows 10 Pro
CPU	Intel(R) i5-10400
RAM	32 GB
Python Version	version 3.8.5

```
8   options = webdriver.ChromeOptions()
9   options.add_argument('headless')
10  options.add_argument('--disable-gpu')
11  options.add_argument('Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/85.0.4183.121 Safari/537.36')
12  driver = webdriver.Chrome('chromedriver.exe',chrome_options=options)
13  driver.implicitly_wait(3)
14
15  driver.get('Piracy Site URL')
16  html = driver.page_source
17  soup = BeautifulSoup(html, 'html.parser')
18  notices = soup.select('specific tag')
19  internal_URL =[]
20  internal_URL = re.findall(r'href=[\'"]?([^\'" >]+)',str(notices))
21  driver.quit()
22
23  link = []
24  for real in internal_URL:
25          html = requests.get(real)
26          soup=BeautifulSoup(html.text, 'html.parser')
27          notices = soup.select('specific tag')
28          link.append(re.findall('href=[\'"]?([^\'" >]+)', str(notices)))
29
30  link = list(filter(None,link))
31
32  f= open("webtoon.csv","w", encoding="UTF-8")
33  for i in range(len(link)):
34      f.write("\n".join((str)link[i]))
35  f.close()
```

Python Selenium Driver

Process of Extracting Internal URL

Process of Extracting Piracy Site

Create Data Set of Piracy Site

Fig. 4. Proposed algorithm overview

❒ Python Selenium Driver

In implementing this algorithm, abnormal traffic is sometimes prevented at the link collection site, so the python selenium module was used to solve this problem.

❒ Process of Extracting Internal URL

In extracting the internal URL, we used the features that exist in the form of 'a *href*' within a specific tag.

❒ Process of Extracting Piracy Site

In extracting the piracy site URL, the piracy site was extracted by accessing the inside of the internal URL and using the features of the URL within a specific tag.

❒ Create Data Set of Piracy Site

In this process, the extracted piracy site was saved in csv format as a database.

The results using the proposed scheme in chapter 4 are shown in **Table 2**, and the proposed scheme has been used to derive 84 torrent sites, 73 video streaming sites, and 79 webtoon sites that infringe copyright.

Table 2. Result of the proposed scheme

Category	Number of Detected Sites
Torrent	84
Video Streaming	73
Webtoon	79
Total	236

However, even if piracy sites are derived from illegally operated link collection sites, some limitations include normally posting sites, not piracy sites, and blocked sites. It is also necessary to refine the proposed scheme, such as the presence of several same URLs within the internal URL.

6. Conclusion

In this paper, we proposed an architecture for automatically detecting continuously appearing piracy sites to prevent copyright infringement due to increased copyright market size and diversification of content. Based on the proposed detection techniques, the experiment results were also derived from the Data Set. The detection architecture of piracy sites proposed in this paper can respond to continuously appearing piracy sites.

References

[1] D.H. Kim, H.S. Jeong, and J. Kwak, “Development of Lifecycle Model for Copyright Infringement Site”, *Journal of The Korea Institute of Information Security & Cryptology*, Vol. 30, No. 1, pp.101-121, Feb. 2020.

[2] Korea Manwha Contents Agency, “Southeast Asia Cartoon Market Settlement”, *Overseas Cartoon Industry Report*, p.5, Dec. 2018.

[3] E.J. Kim, S.K. Choi, and J. Kwak, “A Detection of Webtoon Infringement Site Using Page Source Analysis”, *KSII The 15th Asia Pacific International Conference on Information Science and Technology (APIC-IST) 2020*, Jul. 2020.

[4] Korea Creative Content Agency, “A Survey on the illegal Distribution of Cartoon Webtoon”, *KOCCA 18-27*, p.43, Dec. 2018.

[5] J.H. Sa and S.J. Lee, “Real-time Phishing Site Detection Method”, *Journal of the Korea Institute of Information Security & Cryptology*, Vol. 22, No.4, pp.819-825, Aug. 2012.

[6] Y.G. Jung, “A Study on Blocking Criminal Websites Based on Internet Domain Registration Data of Phishing Sites”, *Journal of Digital Forensics*, Vol. 9, No.2, pp.79-96, Dec. 2015.

Usability and Performance Evaluation of Medical Institution Websites in Malaysia

Nor Azman Ismail, Mohd Ainal Farhan Mohamad Johari, Mohammad Ridzwan Syah Irwan, Mohamad Faiz Hakimi Ishak and Su Elya Namira Mohamed
School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia (UTM),
Johor Bahru, Malaysia
[e-mail: azman@utm.my, mohd.ainalfarhan7@gmail.com, ridzwansyahirwan@gmail.com, faizhakimi91@gmail.com, suelyanamira@gmail.com]
*Corresponding author: Nor Azman Ismail

Abstract

Usability of a website has been one of the essential factors in order to differentiate a good and a bad website. Each website must be able to give maximum interaction and satisfaction for its user. A medical institution website must have top-notch usability since the users are mostly will be a patient that lack knowledge or not well informed about health. It will make the users, or the patient can get a good educational content easily for their health care. This study analyses the website and evaluates the usability of 10 medical institutions' websites in Malaysia. Two different methods are used in this study, which is a questionnaire-based evaluation method and usability testing tool method. From this study, we concluded that most of the medical institution websites are lack of usability for users. Therefore, the users will have a hard time to get the satisfaction that they wanted from the service offered on the website. In order to make sure that the websites meet user expectations, the person that's responsible with medical institution websites must update their websites according to usability evaluation and web performance metrics.

***Keywords*:** Hospital website; Usability; Web page analysis; Website evaluation; Healthcare

1. Introduction

Nowadays, the growth of technology and the internet have made the life of many people easier. It is also affecting the number of people that are going to meet doctors at the medical institution since various websites, and social media have its educational content that able to teach people about health care [1]. Medical service is one of the essential services that every citizen need [2]. For example, in Wuhan, China, the epicentre of the novel coronavirus outbreak has recorded nearly 400 doctors providing free consultation services related to coronavirus and disease symptoms on the Tongji hospital's website for free [3].

In 2018, more than 90 per cent of people from 18 to 24 years old said that they would trust medical information shared by others on their social media network and more than two times as likely than 45 to 54 years old use social media for health-related discussions [4]. This shows how important each medical institution appoint their staff to create and manage its websites and health care services to make citizen be able to get authentic information on any type of healthcare [5,6]. According to data collected by Malaysia Statistic Research Department, there were around 144 public medical institutions and 240 private medical institutions in Malaysia in 2018 [7]. The data show that if each public and private medical institutions in Malaysia has its websites,

the variety of information about health care that citizen about can get is wide.

This paper conducts usability evaluation [8] and web performance on official public and private medical institutions in Malaysia. In brief, the objective of this study is: (i) To discover the usability of medical institution websites in Malaysia; (ii) To analyses and compare which medical institution websites fulfilled the website usability evaluation [9]; (iii) To make a suggestion based on the findings on how to make a better website for the medical institution in Malaysia. The next section, we discussed the methodology used in this study, then we show the result of our evaluation findings of selected medical, institutional websites. Finally, the discussion was based on the findings and concluded the overall study.

2. Methodology

In this study, we are using two evaluation methods. The first one is a questionnaire, and the second is based on usability testing tools GTMetrix [10] and Website Grader [11]. This study was involved in ten medical institution websites in Malaysia which listed in **Table 1**.

Table 1. List of selected Malaysia medical institution websites

No.	Medical institution Name
1.	Hospital Kuala Lumpur
2.	Pantai Hospital Malaysia
3.	Sunway Medical Centre
4.	Beacon Hospital
5.	Tung Shin Hospital
6.	Mahkota Medical Centre
7.	Manipal Hospital
8.	Prince Court Medical Centre
9.	KPJ Ampang Puteri
10.	Thomson Hospital

2.1 Questionnaire-based Evaluation Method

In order to evaluate the usability of the websites, the usability aspects of ten medical institution websites in Malaysia were identified and analyzed [12]. From the analysis, the relevant question for the Questionnaire-based Evaluation Method has been developed based on the following five key factors [13] (**Table 2**). A total of fifty respondents has participated in our survey to evaluate selected medical institution websites in Malaysia.

Table 2. Five keys factor

No.	Key Factor
1	Content, organization, and readability
2	Navigation and links
3	User interface design
4	Performance and effectiveness
5	Educational information

The questionnaire consists of two sections. The first section consists of the respondent information, which includes name, age, their experience surfing one of the medical institution websites in Malaysia, and how often did they access with the medical institution websites. The second section consists of the five key factors listed in **Table 2**.

2.2 Usability Testing Tools Method

The second method is by using the web performance testing tools, which is Website Grader and GTMetrix to test the usability of all the selected websites. Website Grader is the tools for measuring the effectiveness of a website and provides an inbound marketing score while GTMetrix is a tool that analyses the websites page's speed performance. The test consists of four main factors which are performance, number of requests of the websites, load time, and page size. These testing methods are used to observed how fast the websites will process each information from the user and the ability to perform many tasks at one time. It is to ensure that the websites will be able to manage many users at one time without error.

3. Result and Discussion

3.1 Results of Questionnaire-based Evaluation Method

From the questionnaire given, statistically, we found that most of our respondent age is around 19 to 25 years old. For the gender, we manage to get 16 respondents were female, and the rest, which is 34 respondents were male. The majority of the respondents we got to say that they often visit the medical institution website to find new information about health care. 76% of the respondents visit the website regularly, and the rest rarely visited the website.

For the rest of the questionnaire, results are as shown in **Table 3** below. It has shown that some of the respondents are dissatisfied with the performance of the medical institution websites in Malaysia. In general, all selected websites have useful educational information since they got the highest value which is 48% on strongly agree with it. Then followed by the navigation and links key factor where this factor got 24% strongly agree from our respondent. For the rest of the key factors, user interface design, performance and effectiveness and content, organization and readability they got 12%, 12%, and 14%. Key factor content, organization, and readability got the most value for the strongly dissatisfied. From there, we can see that there was a huge gap from one key factor compare to another key factor. It shows that the websites still need improvement in other to meet the user requirement of the websites.

3.2 Performance

Page Speed and YSlow are an indication for the front-end of a website best practices for optimal speed [14]. Based on **Fig. 1**, princecourt.com is the best in terms of its Page Speed and YSlow score. However, it is not the website which has the most excellent performance compared to others. Four websites have a more outstanding performance with thomsonhospital.com as the first place with the performance of 63.33%. According to GTmetrix (2017), front- end and back-end of a website are the one who determines the performance of a website [14]. Even though princecourt.com has an advantage in term of Page Speed and YSlow score, the performance is low than expected due to the inefficiency of the back end of the website.

3.3 Load Time

Based on **Fig. 2**, there is a 13.4 seconds difference in the load time of sunwaymedical.com when comparing the result of two different testing tools. This drastic difference in results shows the importance of using multiple testing tools in testing a website to enhance the accuracy of the result. Hence, sunwaymedical.com is the website that has the most loading time compare to others with 33 seconds load time; meanwhile, kpjampang.com has the least loading time, which is 10.43 seconds.

3.4 Size of The Page

Based on Figure 3, the increase of a page will lead to the increasing in loading time and eventually make the performance decrease. Based on the two testing tools, princecourt.com has the least size of a page in total, which is 1.66 Megabytes. It is one of the contributors for making the website has more excellent performance compared to almost half of the websites tested. 21.7 Megabytes in a total of its page size, sunwaymedical.com become the website which has the most significant page size. The importance of reducing unnecessary information which may become the burden of the website is compulsory in order to increase the overall performance of the website.

3.5 Number of Request

Based on **Fig. 4**, the number of requests per website is as essential as the others because the increase requests in a website will lead to an extended period. Most of the websites tested only has around 48 to 100 requests, but sunwaymedical.com has many requests that need to be taken care of with 237 requests from gtmetrix.com and 204 requests from using website.grader.com. This one of the reasons that make a website having a bad performance.

Table 3. Evaluation of websites by the user

Features	Usability of websites by the user (%)				
	Strongly Dissatisfied	Dissatisfied	Fair	Agree	Strongly Agree
Content, organization, and readability	40	12	22	12	14
Navigation and links	8	12	36	20	24
User interface design	10	16	20	42	12
Performance and effectiveness	28	32	24	10	12
Educational information	8	12	10	22	48
Total	**18.8**	**16.8**	**22.4**	**20**	**22**

Table 4. Result from GTMetrix

No.	Hospital website	Elements to be measured				
		Page Speed Score (%)	YSlow Score (%)	Fully Loaded Time (s)	Total Page Size (MB)	Request
1.	hkl.gov.my	27.0	48	8.50	6.32	77
2.	pantai.com.my	40.0	69	7.00	2.84	69
3.	sunwaymedical.com	24.0	59	23.20	12.0	237
4.	beaconhospital.com.my	58.0	61	8.60	3.29	95
5.	tungshin.com.my	25.0	58	3.90	3.49	85
6.	mahkotamedical.com	67.0	57	5.70	3.04	145
7.	manipalhospitals.com.my	56.0	72	3.90	3.49	100
8.	princecourt.com	95.0	74	7.80	0.96	63
9.	kpjampang.com	72.0	69	4.13	1.49	93
10.	thomsonhospitals.com	81.0	73	4.80	1.42	52
	Average	**54.5**	**64**	**7.75**	**3.91**	**101.6**

Table 5. The result from Website Grader

No.	Hospital website	Elements to be measured			
		Performance (%)	Page Size (MB)	Page Request	Page Speed (s)
1	hkl.gov.my	40.00	0.27	73	10.50
2	pantai.com.my	50.00	1.60	70	3.80
3	sunwaymedical.com	23.33	8.90	204	9.80
4	beaconhospital.com.my	30.00	1.60	94	10.90
5	tungshin.com.my	33.33	0.59	83	7.10
6	mahkotamedical.com	23.33	6.80	162	9.60
7	manipalhospitals.com.my	56.67	1.70	97	6.80
8	princecourt.com	46.67	0.66	61	9.50
9	kpjampang.com	56.67	0.99	91	6.30
10	thomsonhospitals.com	63.33	0.69	48	6.90
	Average	**42.33**	**2.38**	**98.3**	**8.12**

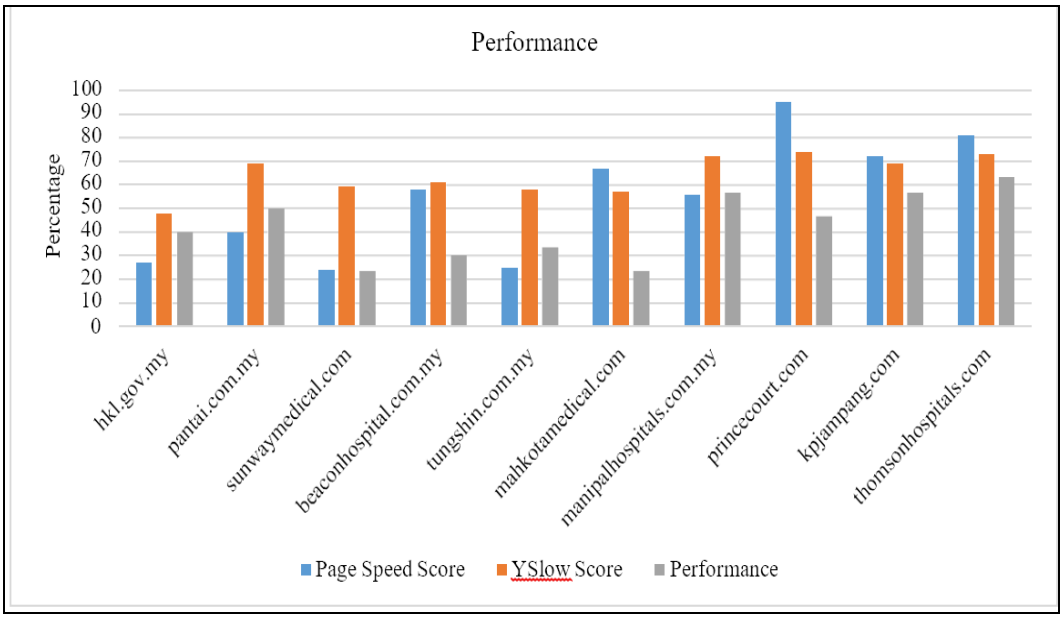


Fig 1. Performance of the selected websites

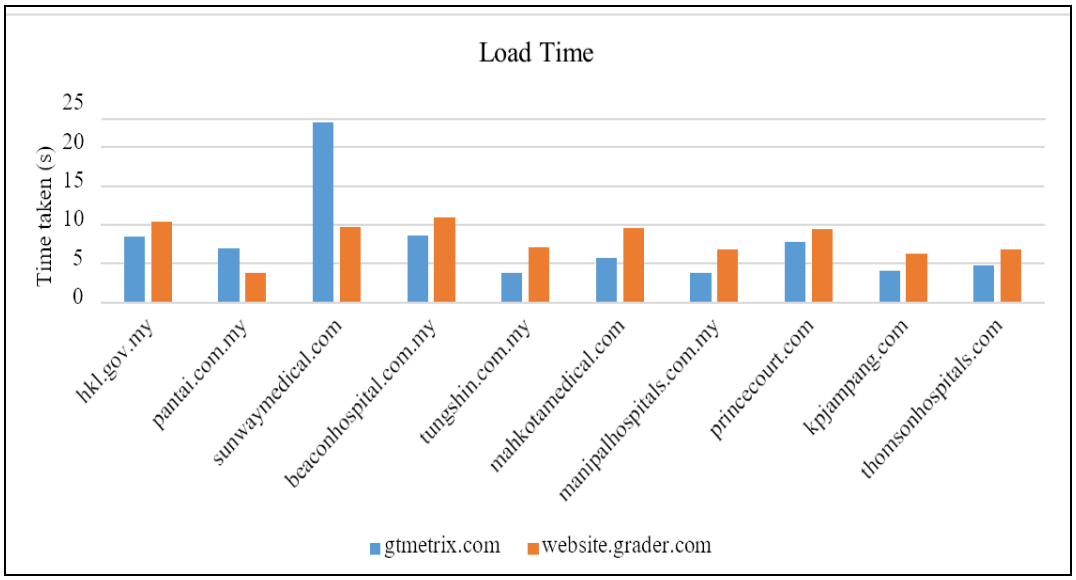


Fig. 2. Load time of the selected websites

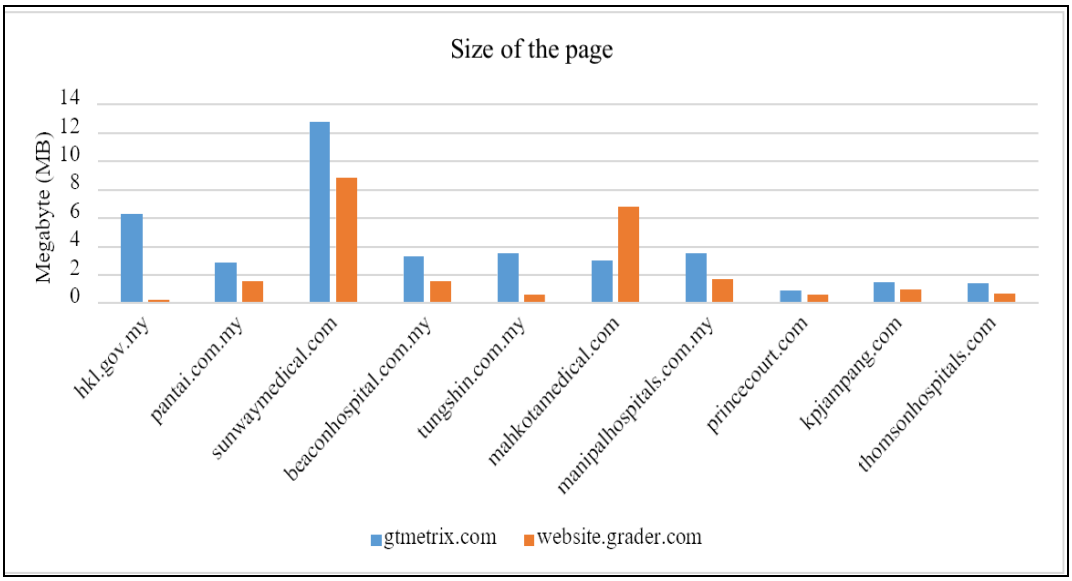


Fig. 3. Size of the page

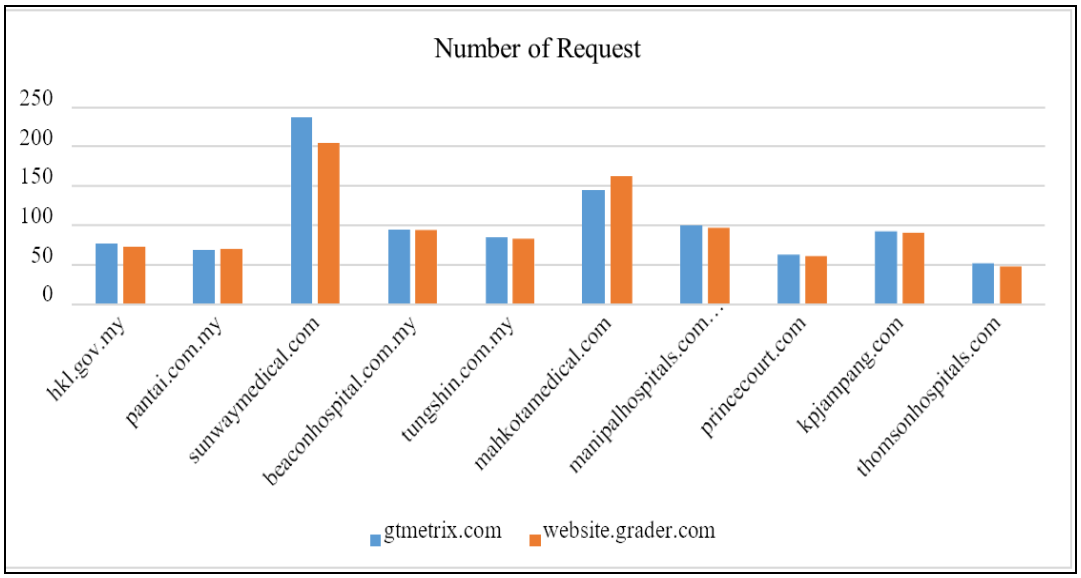


Fig. 4. Number of requests of the selected website

4. Conclusions

Based on the present study, it showed that the features of the medical institution websites in Malaysia have limited features of the medical, institutional website. This leads to failure of meeting the end-user demands and expectation towards the medical institution websites. Simultaneously, the website's testing tools reveal that they need to improve in terms of the size of the page, load time, and the number of requests in order to speed up the overall performance.

In assessing the usability of these websites, some key issues need to be taken care of to meet user demand and expectations. The use of more testing tools to diagnose websites can help medical institution webmasters know what to do to improve overall performance on the website. Top three websites are ranked with thomsonhospitals.com in the first place, followed by kpjampang.com and manipalhospitals.com.my based on their overall performance. It will be better to carry out a more comprehensive study which covers more medical institution websites and use more testing tools to increase the efficiency of the result from the study.

References

[1] ReferralMD, 24 Outstanding Statistics & Figures on How Social Media has Impacted the Health Care Industry, 2013, https://getreferralmd.com (accessed on 25 March 2020)

[2] Ittipong K., Watsawee S., & Manachai T., Telemedicine- Meaning, Challenges and Opportunities, Siriraj Medical Journal, Mahidol University, 2019, 71(3), 246-252.

[3] Huaxia, China Focus: Telemedicine reduces hospital pressure, cross infection amid virus outbreak 2020, http://www.xinhuanet.com/english/2020-02/06/c_138760566.htm (accessed on 25 March 2020).

[4] Paul, C., Online Health Information: Is It Reliable?, 2019 https://glhf.org/online-health-information-is-it-reliable/ (accessed on 28 March 2020).

[5] Edgar, H., Clark, K., Gordon, W., Teresa, F., How wired are U.S. hospitals? A study of patient-oriented interactive tools, International Journal of Healthcare Management, Taylor & Francis., 2020, 13(1), 72-78.

[6] Usher W, Skinner J. Health Websites and Reliability Components. ACHPER Australia Healthy Lifestyles Journal. 2007, 55(4), 29-34.

[7] J. Müller, Statista Research Department: Number of public and private hospitals in Malaysia from 2013-2018, 2020. statista.com/statistics/794860/number-of-public-and-private-hospitals-malaysia/ (accessed on 20 March 2020).

[8] Nielsen, J. Heuristic evaluation. In J. Nielsen, & R. L. Mack (Eds.), Usability inspection methods. New York: Wiley.,1994, 25– 62

[9] Akoglu, C. Usability evaluation: A method for a specific field. Department of Communication Design, Istanbul, Turkey: Yildiz Technical University, 2002.

[10] GTmetrix: See how your site performs, reveal why it's slow and discover optimization opportunities. http://www.gtmetrix.com. (accessed on 5 March 2020).

[11] HubSpot, Website Grader: How strong is your website?, 2007, http://www.websitegrader.com. (accessed on 5 March 2020)

[12] Nielsen, J. Finding usability problems through heuristic evaluation.Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: ACM, 1992, 373–380

[13] Battleson, B., & Booth, A. Usability testing of an academic library web site: A case study. J. Acad. Librarianship, 2001

[14] GTmetrix Performance Blog, PageSpeed and YSlow are Half the Battle. 2017, https://gtmetrix.com/blog/pagespeed-and-yslow- are-half-the-battle/ (accessed on 20 March 2020)

Exploring Levels of Credit Risk in P2P Lending: A Quantitative and Qualitative Emotion Mining Method

Min Gyeong Kim[1], Francis Joseph Costello[1], Tolchinski Rita[1], and Kun Chang Lee[1,2,*]
[1]SKK Business School, Sungkyunkwan University
Seoul, South Korea
[2]Professor, Samsung Advanced Institute for Health Sciences & Technology (SAIHST)
Sungkyunkwan University
Seoul 03063, South Korea
[e-mail: mgkim9541@gmail.com, joe.costehello@gmail.com, ritetol@gmail.com, kunchanglee@gmail.com]
*Corresponding author

Abstract

Although peer-to-peer (P2P) loans are growing at a rapid rate of growth, there is a credit risk of information asymmetry that can create a barrier. Many studies have conducted predictive studies of credit risk using numerical and categorical data. However, such a new study using unstructured data, such as text data, is lacking. The text that the borrower uses for the loan application has a great potential to be a useful source of information. We propose a novel credit risk analysis framework based on text mining that can extract emotion and sentiment from borrowers' text. On top of this, we adopt various AI techniques as to identify the most effective model for prediction. Specifically, we look at how the loan description of P2P lending affects loan performance. In doing so we expect to show that our proposed integrated analysis of structured and unstructured data can help to predict credit risk. Furthermore, most prior work has discussed binary outcomes of risk, however, people's tolerance for risk is not binary and has some fuzziness. Thus, we propose a multiclass approach to incorporate different levels of risk. The results of the study reveal how the text of borrowers and the content of the description relate to future credit risk.

***Keywords*:** P2P Lending, Emotion Mining, Regression Analysis, Machine Learning, Multiclass Classification

1. Introduction

The success of a mutual transaction is often threatened by the problem of moral hazard in which a party in a transaction can act behind ones' back. Moral hazard is often more serious when information asymmetry occurs between both stakeholders. For example, moral hazard in e-commerce and the shared economy has been seen to have numerous of implications on trust. In these two settings mutual trust is essential, and thus, research has been actively pursued in this area. Measures to address moral hazard and information asymmetry have shown that traditional solutions (i.e., incomplete monitoring and forced transaction costs) are not only costly, but also inefficient [1,2,3].

This study is based on data from P2P (Peer to-peer) loans, which characterize a financial service in the shared economy or P2P economy. P2P Lending is a disruptive innovation in the financial industry and is booming because they are a convenient means of solving the financial needs of small businesses and individuals.

Our study tries to investigate the role of emotions in borrower self-description, which are shown to prospective creditors. Emotions, when positive or negative, may impact the credit risk assessment of lenders, thus affecting their decision towards approving a loan. We would like to see if the emotion revealed in the loan description is related to credit risk. In doing so, we try to explain the drivers of credit risk by using linear regression models and then use machine learning techniques to confirm that emotions help to predict credit risk.

2. Related Work

Choosing a person to trust is not an easy task. However, making an accurate decision when selecting borrowers who can repay their loan is an even more challenging task [4]. Therefore, based on previous pieces of literature, much effort has gone to investigate the development, evaluation, and application of predictive decision support systems in P2P lending [5].

2.1 Peer-to-peer lending

In this study, we focus on the altruistic aspects of P2P lending, ones that do not explicitly consider monetary benefits. Given that previous studies have widely examined key loan features that influence the success of P2P lending, we seek to investigate the role that unstructured textual information may have on the ability for one to examine P2P lending riskiness.

2.2 Machine learning and credit risk assessment

Previous research on P2P lending credit risk assessment have focused on the use of single classification models such as Neural Networks (NN) [6], Support Vector Machines (SVM) [7,8,9] and k-Nearest Neighbor (kNN) [10]. Attention then turned to the use of ensemble techniques such as Random Forest (RF), Bagging, Random Subspace [11] and Stacking [12]. Recently, Deep learning approaches have been explored [13]. Till now, prior research has agreed that the use of models such as RF are superior in P2P lending research [4].

2.3 Emotion and text mining in P2P

Emotion mining represents a task in which different emotional states are extracted from text [14]. In previous studies, emotion mining has been found to be an effective approach in text mining, especially in the case when readers expose information within the text. Supavich and Frederick found that borrowers can improve their chance to receive a loan if they improve their loan descriptions through the expression of specific emotions [15]. Also, it has been emphasized that emotions, such as trust, could be critical in P2P lending and contribute to borrowers receiving a loan [16]. For these reasons, attempting to understand their effects in P2P lending, and the potential this has for identifying risky loans is a valuable research topic.

2.4 P2P loan descriptions: Emotion in textual information

Extant literature in e-commerce suggests that positive emotions such as satisfaction, trust, and dependence, can be a positive influence on success [17]. Nonetheless, emotions in P2P lending are unexplored. As it has not been revealed how emotions work in P2P lending, it is unclear whether they can be successful in identifying a safe loan. This is because successful funding opportunities can be increased on the basis of lenders' assumptions that are worth worrying about [18]. For example, a lender may be more willing to support a loan after feeling sympathetic to the borrower's difficulties shown in his loan description. As a result, a loan in which grief is the main emotion in the description can potentially invigorate the lender's lending behavior compared with positive emotions.

According to Pang and Lee [25], emotional analysis refers to identification and extraction of subjective information from source materials by using data mining techniques. Literature on emotional analysis has been applied to a number of research areas, including customer reviews [22], blogs [23], social media [24], and crowdfunding [20].

In this study, a series of basic emotions that are thought to have a distinct function in individual motivation and behavior are investigated [19, 20]. Emotions can induce different motivational characteristics [21], and hence we hypothesize:

RQ1: The emotions of the loan description will be associated with a credit risk.

RQ2: The emotion in the description will improve the credit risk forecast performance.

3. Methodology

3.1. Data collection

The dataset is based on a public dataset collected from www.LendingClub.com. Lending Club is a P2P Lending Company that provides financial and loan services in the United States. The company's lending model is based on a typical crowdfunding model, in which any individual can contribute individually to loans or fund certain loans as a member of the lending group. This data set includes a set of variables, including information about borrowers, lenders, loan groups and borrowers.

3.2. Procedure for data preparation

The original data set consisted of P2P loans recovered between 2007 and the third quarter of 2019. However, loans that did not have a description were excluded. The data after preprocessing included 125,819 loans. We examined the delinquencies of the last two years of the loan period to measure credit risk. The following is a table of data classifications based on the delinquencies of the collected data.

As shown in **Table 1**, most loans have not been delinquent. However, some loans resulted in arrears, and 22 were found that caused the most arrears. Delays in paying interest increases credit risk. In this study, we classified credit risk as high if three or more arrears occurred, just as our credit rating would fall if interest on a loan carried out by a bank was not paid normally.

Table. 1. Data Distribution

Delinquency	*Number*	*Credit Risk*
0	106,190	Low Risk
1 ~ 2	16,517	Medium Risk
Over 3	1,936	High Risk

3.4. Classification and Sampling Analysis

Since the data has an unbalanced distribution, classification performance was weak and needed addressing. A common technique known as sampling was used. The purpose of the sampling method is to ensure that the distribution of classes in a data set reaches a balanced distribution. Sample methods include under-sampling, over-sampling, over-under sampling, and specialist sampling such as SMOTE. The under-sampling method removes several examples from the main class. Oversamples, on the other hand, are the re-sampling or generation of minority cases. Another approach that draws attention in oversampling is synthetic data generation. SMOTE (Synthetic Mineral Over-Sampling Technology) is one of the best-known methods. All approaches were analyzed.

4. Analysis and results

4.1. Variables Description

Variables selected for analysis are seen in the following:

- **Emotion:** Anger, Anticipation, Disgust, Fear, Joy, Sadness, Surprise, Trust, Negative, Positive.
- **Quantitative Variables: A**nnual income, credit age, delinquencies (recent two years), employment length, home ownership, open accounts (current credit lines), total accounts (total num. of credit lines), term (length of loan).
- **Financials:** DTI (debt to income ratio), income to payment ratio, revolving utilization rate, revolving to income ratio, FICO Score, credit decisions indicator.
- **Nominal variables:** Loan status (class attribute), grade (given by Lending Club).

The dependant variable in the regression analysis was Credit RiskOLS which contains High Risk: 2, Medium Risk: 1, Low Risk: 0; Logistic: High and Medium Risk: 1, Low Risk:0).

4.2. Regression model

In the regression model, the ratings on the core emotion such as sadness ($p < 0.01$), negative ($p< 0.05$) and positive ($p<0.1$) were seen to be significant. The analysis of low credit risk from loan description revealed 'Sadness', a negative emotion, as significant, consistent in both analyses. The sad message can be interpreted as a factor which potnetially reduces credit risk. Interestingly, the logistic regression shows that 'disgust' increases credit risk ($p < 0.1$).

Table. 2. Descriptive Statistic

	Mean	*SD*	*Max*	*min*
anger	0.19	0.54	11	0
anticipation	1.39	1.98	44	0

disgust	0.05	0.26	6	0
fear	0.21	0.62	22	0
joy	1.15	1.62	44	0
sadness	0.77	1.15	21	0
surprise	0.32	0.77	19	0
trust	2.30	2.67	55	0
negative	2.32	1.90	40	0
positive	3.03	3.51	77	0
int_rate	13.45	4.28	26.06	5.42
installment	425.70	239.74	1408.13	4.93
annual_inc	72266.57	52988.38	7141778	1896
dti	16.34	7.49	38.57	0
delinq_2yrs	0.22	0.67	22	0
CR_age	5.60	0.28	6.732	4.762
open_acc	10.78	4.67	62	1
total_acc	24.46	11.33	105	1
IPR	-2.68	0.61	-0.898	-8.149
revol_util	56.20	24.23	120.2	0.01
RIR	-4.70	0.88	-0.589	-13.942
FICO	702.97	32.46	847.5	612

4.3. Classification

Tthe original dataset results are very poor due to the unbalanced dataset. Although the accuracy is high, as can seen within the high risk (HR) true positive rate, the prediction is extremely low. After sampling, the results how that RF is undoubtedly the best model in all of the sampling techniques. Precisely, the RF showed accuracy of 94.27%, 94.30%, 94.31%, and 93.03% in the over, under, over-under, and SMOTE sampling respectively. The most superior performance was seen in the Under-Over sampling combined with RF technique whereby the AUC (.992), true positive rate for MR (.923) and HR (.997) coupled with accuracy make it ideal for predicting the various levels of credit risk.

5. Conclusions

P2P financing can provide financial support to individuals who have difficulty using financial services. But our understanding of the different P2P loan characteristics and the lender's credit risk is limited. It is not clear how those features predict credit risk, given the various information and the rich features inherent in P2P loans.

Based on RQ1, our results indicate that emotion in text, i.e., "Sadness" shows lower credit risk. Judging from these results, it can be seen that the feelings of a loan description are related to credit risk.

For RQ2, We have analyzed the data with classification models. In Original Data, accuracy figures are high, but in practice they have failed to predict credit risk (**Table 4**). The sampling technique with the best numerical performance is Under-Over Sampling (SMOTEENN). Of the four classifiers, RF showed the best performance.

Table. 3. Regression Results

	OLS		*Logistic*	
	Estimate	***Std. Error***	***Estimate***	***Std. Error***
ADj	0.104			
AIC			90673.060	
(Intercept)	*3.987****	*0.082*	*17.066****	*0.405*
anger	*0.007*	*0.005*	*0.038*	*0.022*
anticipation	*-0.002*	*0.002*	*-0.007*	*0.012*
disgust	*0.011*	*0.008*	*0.064**	*0.038*
fear	*0.002*	*0.004*	*0.016*	*0.018*
joy	*0.002*	*0.003*	*0.009*	*0.015*
sadness	*-0.008****	*0.003*	*-0.036****	*0.014*
surprise	*0.005*	*0.004*	*0.019*	*0.018*
trust	*-0.002*	*0.002*	*-0.008*	*0.010*
negative	*0.004***	*0.002*	*0.015**	*0.009*
positive	*-0.003**	*0.002*	*-0.014**	*0.008*
term60 months	*-0.007*	*0.006*	*0.029*	*0.025*
int_rate	*0.008****	*0.002*	*0.034****	*0.006*
installment	*1e-04**	*0.000*	*5e-04****	*0.000*
gradeB	*-0.034****	*0.008*	*-0.098***	*0.039*
gradeC	*-0.057****	*0.012*	*-0.252****	*0.054*
gradeD	*-0.064****	*0.016*	*-0.339****	*0.069*
gradeE	*-0.083****	*0.019*	*-0.481****	*0.084*
gradeF	*-0.084**	*0.024*	*-0.568****	*0.101*
gradeG	*-0.033*	*0.031*	*-0.527****	*0.122*
home_own NONE	*0.164*	*0.123*	*0.564*	*0.433*
home_own OTHER	*0.055*	*0.054*	*0.236*	*0.225*
home_own OWN	*-0.036****	*0.007*	*-0.167****	*0.033*
home_own RENT	*-0.057****	*0.004*	*-0.276****	*0.019*
annual_inc	*0.000*	*0.000*	*0.000***	*0.000*
dti	*0.000*	*0.000*	*0.003**	*0.001*
CR_age	*0.271****	*0.008*	*1.074****	*0.033*
open_acc	*-0.014****	*0.001*	*-0.044****	*0.003*
total_acc	*0.007****	*0.000*	*0.024****	*0.001*
IPR	*-0.034****	*0.006*	*-0.183****	*0.029*
revol_util	*-0.004****	*0.000*	*-0.016****	*0.001*
RIR	*-0.073****	*0.005*	*-0.326****	*0.023*
FICO	*-0.008****	*0.000*	*-0.038****	*0.001*

□□□ $p < 0.01$. □□ $p < 0.05$. □ $p < 0.01$.

Table. 4 Prediction results of risky P2P loans based on sampling and classification

Sampling	Model	Acc	AUC	True Positive Rate			False Positive Rate		
				NR	MR	HR	NR	MR	HR
Original Data	Logistic	84.86%	.765	.991	.033	.000	.960	.011	.000
	kNN	76.93%	.554	.870	.208	.049	.757	.121	.014
	RF	85.1672	0.752	0.999	0.003	0.000	0.996	0.001	0.000
	NN	85.19%	.756	1.000	.001	.000	.999	.000	.000
Over sampling	Logistic	69.88%	.866	.878	.480	.739	.124	.165	.163
	kNN	90.32%	.937	.842	**.870**	**.995**	**.052**	.070	.023
	RF	**94.27%**	**.988**	**.983**	.850	**.995**	.076	**.009**	**.002**
	NN	77.22%	.901	.900	.598	.818	.098	.108	.136
Under sampling	Logistic	69.90%	.866	.879	.480	.739	.124	.165	.163
	kNN	90.36%	.937	.845	**.870**	**.995**	**.052**	.070	.023
	RF	**94.30%**	**.988**	**.984**	.851	**.995**	.076	**.008**	**.001**
	NN	77.15%	.900	.903	.591	.820	.100	.104	.139
Under-Over Sampling (SMOTEENN)	Logistic	70.64%	.862	.809	.554	.765	.068	.180	.200
	kNN	90.61%	.941	.812	.890	**.997**	.042	.077	.022
	RF	**94.31%**	**.992**	**.899**	**.923**	**.997**	**.036**	**.045**	**.002**
	NN	76.83%	.898	.816	.594	.892	.059	.106	.189
SMOTE	Logistic	69.29%	.858	.965	.440	.674	.140	.133	.188
	kNN	88.56%	.914	.831	**.840**	**.986**	**.052**	.077	.043
	RF	**93.03%**	**.985**	**.980**	.832	.979	.080	**.014**	**.011**
	NN	70.46%	.864	.939	.452	.723	.117	.135	.191

References

[1] M. Hallsworth, J.A. List, R.D. Metcalfe, and I. Vlaev. "The Making of Homo Honoratus: From Omission to Commission." *National Bureau of Economic Research*, 2015.

[2] F. Lu, J. Zhang, and J.M. Perloff. "General and Specific Information in Deterring Traffic Violations: Evidence from a Randomized Experiment." *Journal of Economic Behavior & Organization* 123: 97-107. 2016.

[3] D. Chen, F. Lai, and Z. Lin. "A Trust Model for Online Peer-to-Peer Lending: A Lender's Perspective." *Information Technology and Management* 15, no. 4: 239-54. 2014.

[4] M. Malekipirbazari, and V. Aksakalli. "Risk Assessment in Social Lending Via Random Forests." *Expert Systems with Applications* 42, no. 10: 4621-31. 2015.

[5] S. Lessmann, B. Baesens, H.-V. Seow, and L.C. Thomas. "Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring: An Update of Research." *European Journal of Operational Research* 247, no. 1: 124-36. 2015.

[6] K.K. Lai, L. Yu, S. Wang, and L. Zhou. "Credit Risk Analysis Using a Reliability-Based Neural Network Ensemble Model." *Paper presented at the International Conference on Artificial Neural Networks*, 2006.\

[7] L. Yu, X. Yao, S. Wang, and K.K. Lai. "Credit Risk Evaluation Using a Weighted Least Squares Svm Classifier with Design of Experiment for Parameter Selection." *Expert Systems with Applications* 38, no. 12: 15392-99. 2011.

[8] L. Yu, S. Wang, and J. Cao. "A Modified Least Squares Support Vector Machine Classifier with Application to Credit Risk Analysis." *International journal of information technology & decision making* 8, no. 04: 697-710. 2009.

[9] Y. Wang, S. Wang, and K.K. Lai. "A New Fuzzy Support Vector Machine to Evaluate Credit Risk." *IEEE Transactions on Fuzzy Systems* 13, no. 6: 820-31. 2005.

[10] W. Henley, and D.J. Hand. "Ak‐Nearest‐Neighbour Classifier for Assessing Consumer Credit Risk." *Journal of the Royal Statistical Society: Series D (The Statistician)* 45, no. 1: 77-95. 1996.

[11] G. Wang, J. Ma, L. Huang, and K. Xu. "Two Credit Scoring Models Based on Dual Strategy Ensemble Trees." *Knowledge-Based Systems* 26: 61-68. 2012.

[12] K.-S. Hwang, and S.-B. Cho. "Landmark Detection from Mobile Life Log Using a Modular Bayesian Network Model." *Expert Systems with Applications* 36, no. 10: 12065-76. 2009.

[13] C. Luo, D. Wu, and D. Wu. "A Deep Learning Approach for Credit Scoring Using Credit Default Swaps." *Engineering Applications of Artificial Intelligence* 65: 465-70. 2017.

[14] N. Gupta, M. Gilbert, and G.D. Fabbrizio. "Emotion Detection in Email Customer Care." *Computational Intelligence* 29, no. 3: 489-505. 2013

[15] S.F. Pengnate, and F.J. Riggins. "The Role of Emotion in P2P Microfinance Funding: A Sentiment Analysis Approach." *International Journal of Information Management* 54: 102138. 2020.

[16] Chen, Dongyu, Fujun Lai, and Zhangxi Lin. "A trust model for online peer-to-peer lending: a lender's perspective." *Information Technology and Management* 15.4: 239-254. 2014

[17] Y. Hwang, and D.J. Kim. "Customer Self-Service Systems: The Effects of Perceived Web Quality with Service Contents on Enjoyment, Anxiety, and E-Trust." *Decision support systems* 43, no. 3: 746-60. 2007

[18] S. Black. "Microloans and Micronarratives: Sentiment for a Small World." *Public Culture* 21, no. 2: 269-92. 2009.

[19] P. Ekman. "An Argument for Basic Emotions." *Cognition & emotion* 6, no. 3-4: 169-200. 1992.

[20] Izard, Caroll E. "Organizational and motivational functions of discrete emotions.", 1993.

[21] K.R. Scherer. "Emotion Experiences across European Cultures: A Summary Statement." 1986.

[22] C.-S. Yang, C.-H. Chen, and P.-C. Chang. "Harnessing Consumer Reviews for Marketing Intelligence: A Domain-Adapted Sentiment Classification Approach." *Information Systems and e-Business Management* 13, no. 3: 403-19. 2015.

[23] S. Mukherjee, and P.K. Bala. "Gender Classification of Microblog Text Based on Authorial Style." *Information Systems and e-Business Management* 15, no. 1: 117-38. 2017.

[24] S.M. Zavattaro, P.E. French, and S.D. Mohanty. "A Sentiment Analysis of Us Local Government Tweets: The Connection between Tone and Citizen Involvement." *Government information quarterly* 32, no. 3: 333-41. 2015.

[25] B. Pang, and L. Lee. "Opinion Mining and Sentiment Analysis." *Comput. Linguist* 35, no. 2: 311-12. 2009.

Sentiment Analysis: Study of Social Behaviour of Youth in Different Social Class

Nor Azman Ismail, Siti Noorafishah Abd Latif, Su Elya Namira Mohamed
School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia
Skudai, Malaysia
[e-mail: azman@utm.my, sitinoorafishah@gmail.com, senamira2@live.utm.my]
*Corresponding author: Nor Azman Ismail

Abstract

Advance social media have changed the way people communicate, sharing content, interact and communicate. Nowadays, social media are rapidly becoming a platform for sharing knowledge and opinions regarding particular topics and interest. Sentiment analysis is an active research area of natural language processing with different approaches to understanding views on social media data. Sentiment analysis rapidly growth throughout all branches of studies and academic fields. However, despite the rapid growth of research in this area, few attempts have been made to systematically review on sentiment analysis based on social classes group regarding posting behaviour. Social classes group divided into two groups which are B40 and M40 regarding their family gross income. The main objective of this paper is to provide a better understanding of sentiment analysis from Twitter. From on the study, we identified several approaches of sentiment analysis methods to classify Twitter data into sentiments. The technique used is lexicon based on VADER dictionary to categorize data into positive, neutral and negative sentiments. The study indicated that, although SA is increasingly used for gaining opinion, a better understanding of the landscape and direction is not well reported. This study shows the discussion on B40 and M40 towards positive and negative tweets. T-test conducted to analyses the difference between both social classes group. The study contributions focus on analyzing the distinction of posting sentiment on different social classes group.

***Keywords*:** Sentiment analysis, Social media analytic, Comparative analysis

1. Introduction

Over the past decade, social media have been peak interest diverse field of study as it is the most used mediums of exchanging information, communications and opinions. Social media become a necessity in different areas of the industry to increase work performance and capacity. Social media have been proven to be a valuable platform for information and communication between society [1]. Social networking contributes towards user engagement in social media by increasing the interaction rate and expand communication between peers.

One of the most massive contributions of social media users is known as youth. Youth is a group of people between the age of 15 until 24 years old. Youths have a higher tendency to owning a social media account. The engagement of youth in social media derive them from delivering a particular type of behaviour regarding their post. One of the aspects that differentiate the behaviour of posts of youth in social media is the informal learning gain from their parents. Different social classes of a family determine the knowledge expose towards them. The early teaching of parent significantly affects the youth on social engagement and emotional aspect.

This article contributes towards the classification of sentiment analysis approach used in social media mainly Twitter on how to measures the sentiments of youth classified by social classes group towards online posting behaviour. We analyze ten respondents respective towards two social classes group, B40 and M40. Using the respondent surveys and review, allow us to classify their social media data into positive, neutral and negative sentiments regarding their post. Our analysis will focus on classifications, methods and analysis used to benchmark the sentiment analysis approach towards social media data to extract sentiments behind it. Firstly, we discussed how social classes group is classified and the requirements needed to be group into certain social classes. Then, we introduce the reviews and approaches available to conduct sentiment analysis in social media, methods used in constructing the sentiment analysis process and analyze the data gained after the classification of sentiment analysis. Finally, we described the result of sentiment analysis based on lexicon approach towards the posting behaviour of youth regarding their social classes.

2. Literature Review

2.1 A Study on Social Classes Group

Classification of a family starts from the generalization of household income. The mean income of a household determines the class of family referred. The higher amount of household consumption, the higher the level of economic wellbeing. Classification of family differs in each country. The country net values are considered along with the rate of the currency. According to Income and Poverty in the United States (2016), the median of household income can be explained through the type of household, race and Hispanic origin of householder, age of householder, nativity of the householder, region, residence and earning of year-round workers [2].

2.2 Categorization Based on Household Income

Household income can be calculated based on their median monthly salaries. Household income of a family includes all the wages from the breadwinner of the family. According to Khazanah Research Institute (2014), the distribution of household income reflects the individual income distribution [3]. The amount of the salary, wages and bonus will include calculating the monthly payment of a household income. Household income also can be categorized based on median income by the household group. In Malaysia, the household group can be categorized into three different groups which are Top 20 (T20), Middle 40 (M40), Below 40 (B40). The median household income distribution in 2016 for T20 is RM 13148.00, M40 is RM 6275.00, B40 is RM 3000.00 [4].

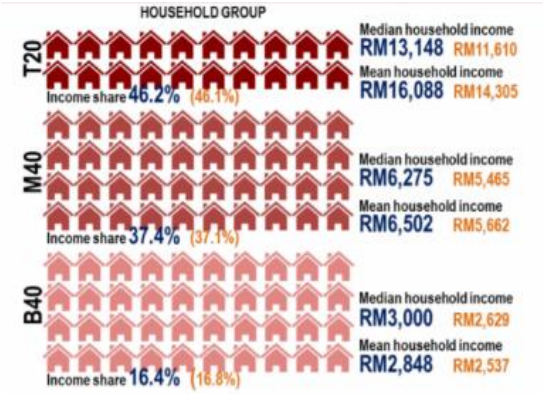


Fig. 1. Distribution of household income group by [4]

2.3 Lexicon Based Approach

Sentiment analysis is known as a process of identifying a combination of words, tone and writing to classify sentiments as positive, negative or neutral. In sentiment analysis, there are two widely approaches used, which are lexicon-based and machine learning. In this section, the lexicon-based approach will be discussed and elaborate. The lexicon-based approach is conducted by assigning a polarity score, which the sum of the polarities will be computed as the overall polarity score of the text [5]. The process counts the number of positive and negative words in text data, and the larger count will be treated as the sentiment of the text. Dictionary approach or also known as lexicon-based approach relies on a lexicon or dictionary of words with pre-calculated polarity [5]. The list of dictionary words is labelled with negative or positive polarities. The dictionary words or lexicons can be created manually.

Lexicon created consider default sentiment words, negation words, blind negation words and split words. The creation process of the lexicon can be conducted in different ways. One of the most commonly used is hand-tagged lexicons. The process involved on manually construct a lexicon by reading several thousands of messages and manually selecting words, which carry sentiments. The method is used to create a lexicon-based approach by establishing a lexicon of words with assigned polarity scores followed by pre-processing and sentiment score calculation [5]. In sentiment calculation, the word is compared with the lexicons. If the word exists in the lexicon, the sentiment score of that word is added to the sum sentiment score of the text. The calculation of the sentiment as follow; 'total sentiment score' = +0.92 +1 +1 +1 -0.84 = 3.08.

$$"A\ masterful[+0.92]\ film\ from\ a\ master[+1]\ filmaker,\ unique[+1] in\ its\ deceptive\ grimness, compelling[+1]\ in\ its\ fatalist[-0.84]\ worldview."$$

Fig. 2. The calculation polarity of a text [5].

2.4 Data Pre-Processing

Most of the social media user used informal language while engaging in social media. Social media user creates their own words and spelling shortcuts and punctuation, misspellings, slang, new words, URLs, and genre specific terminology and abbreviations [6]. Online texts usually contain lots of noise and uninformative parts, such as HTML tags [7]. The ambiguous word used in social media does not have the concrete polarity of sentiments behind it. The noise increases the dimensionality of the process and creates the classification process more difficult [7], which is known as data pre-processing. Data processing involve the process of cleaning ambiguous data to extract meaningful sentiments and polarity of the data. The process of data cleaning is known as text or data normalization. Text normalization involves on analyzing the text HTML characters, slang words, emoticons, stop words, punctuations, URLs' are needed to be removed [6].

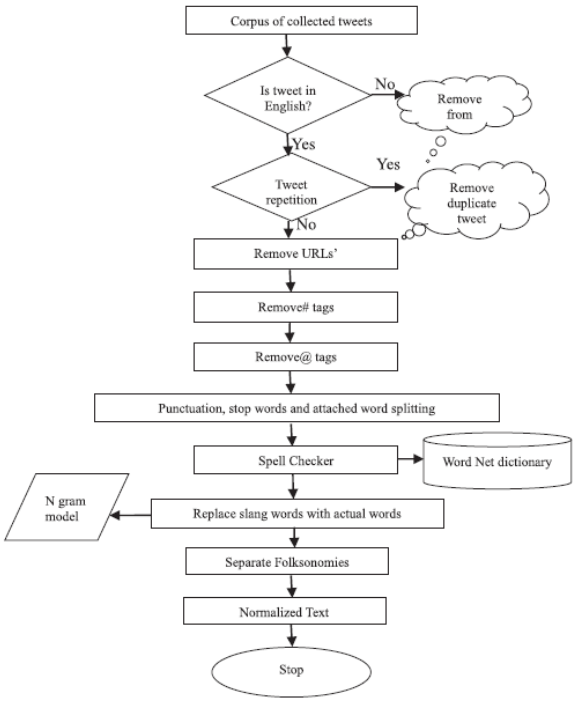


Fig. 3. Text Normalization on Tweets Corpus [6].

Text pre-processing involves polishing and preparing data by removing the removal of punctuation and symbols, tokenization, stemming, and stop word [7]. Process of breaking a stream of text into words, phrases, symbols, or other meaningful elements called tokens are known as tokenization [8]. After the terms are broken down into tokens, stop words removal conducted as they did not contribute towards the content of textual document [8]. The last process involves stemming the words in the textual document. Stemming is the process of conflating the variant forms of a word into a common representation the stem [8].

3. Research Methodology

Fig. 4. Flow of Research Methodology.

3.1 Data Acquisition

The chronology of data acquisition starts with collecting a sample on youth social media classes. Firstly, samples are taken from age 15 until 24 years old randomly and provided with a survey of engagement used in social media. The

participants are given a survey on youth engagement in social media. The survey is fed into two languages which are Bahasa Melayu and English. User social classes are divided into three different groups, Top 20 (T20), Middle 40 (M40), and Below 40 (B40) from monthly household income.

Data crawling using Twitter API are required to gain access on user timeline of tweets and retweets. In this study, Twitter API is used to crawl user tweets information or text by using the username account. The GET statuses/user_timeline used to return a collection of user timeline based on screen_name or user_id parameters. The timeline of the tweets created is parsed for the analysis of the sentiments. The object accepts the parameters of the count, tweet_mode to gain all tweets characters and text. The method can return up to 3200 of user most recent tweets. The tweets will be store into a JSON format that contains created_at and full_text value.

3.2 Twitter Text-Data Pre-Processing

Text data pre-processing is the action of removing noise in-sample data to extract meaningful sentiment on tweets. Data cleaning is the process of cleaning tweets into normalize tweets text. The process requires regular expression to filter the characters in a large number of tweets crawled. The flowchart of data cleaning is discussed as follow.

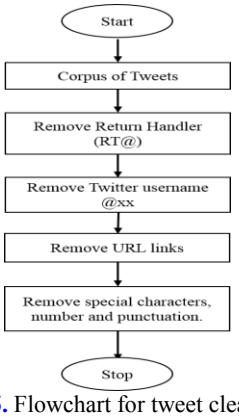


Fig. 5. Flowchart for tweet cleaning

The process starts with preparing corpus of tweets. The corpus of tweets will pass into the process of removing the return handler (RT@). After the return handlers are removed, the twitter username that follows will be removed. Tweets contain links and shortened URLs. The links are not required for sentiment analysis; therefore, the links and URLs are removed. Lastly, special characters, numbers and punctuation will be drawn to gain cleaned compound of words. The clean tweets passed through sentiment analysis analyzer to calculate the score of the compounds.

3.3 Data Analysis

The lexicon dictionary used are Valence Aware Dictionary and Sentiment Reasoner (VADER). Python programming language used as it supports standard package manager (pip) of VADER library. The method and process approach of VADER is as follow:

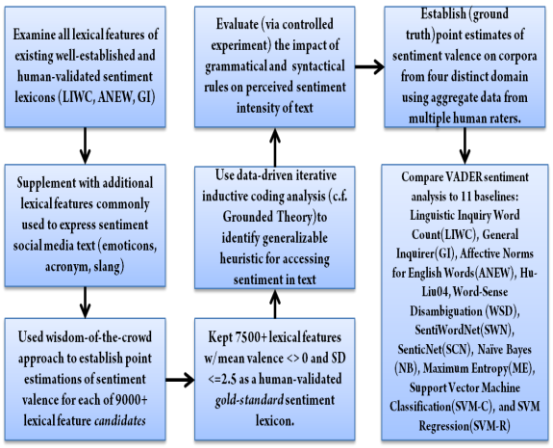


Fig. 6. VADER method and approach process.

After VADER library applied, pip Google Translate install to convert the tweets into English. The standard scoring method of word compound is based on English words. The tweets are converted using translator.translate('tweets').text to English. Next, sentiment analyze analyzer called from VADER library to break down tweets into the scoring compound by using analyser.polarity_scores. Polarities of tweets sentiments are calculate using the analyser.polarity_scores on each tweet and classify into positive, negative and neutral.

4. Result Analysis and Discussion

4.1 Research Result and Analysis

Results show the counts of positive, neutral and negative tweets gained from the sentiment analysis classification. The numbers of tweets

are a map with the social classes group B40 and M40. From the table above, ten respondents are tested, and the number of tweets sentiments are tabulated. To prove the relationship between social classes group and sentiments of tweets, Independent T-test was conducted to find the mean difference of B40 and M40 in positive and negative tweets that map towards their online behaviour posting. A null and alternative hypothesis is created to determine the difference between B40 and M40 sentiments.

Table 1. Flowchart for tweet cleaning

Social Classes Group	Numbers of Tweet Sentiments		
B40	Positive	Neutral	Negative
User 1	555	1031	414
User 2	420	1403	177
User 3	470	1265	265
User 4	843	836	321
User 5	498	1056	446
M40			
User 1	276	1634	90
User 2	364	1500	136
User 3	646	1223	131
User 4	85	62	16
User 5	322	1480	198

From the tabulated table above, the null hypothesis and alternative hypothesis are made.

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

$H_0 = B40 \text{ and } M40 \text{ positive tweets means are equal}$

$H_1 = B40 \text{ and } M40 \text{ positive tweets means are not equal}$

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

$H_0 = B40 \text{ and } M40 \text{ negative tweets means are equal}$

$H_1 = B40 \text{ and } M40 \text{ negative tweets means are not equal}$

Table 2. T-test of social classes group and positive sentiments

	pos_tweets	
Levene's Test for Equality of Variances	**Equal variances assumed**	**Equal variances not assumed**
F	.058	
Sig.	.815	
t-test for Equality of Means		
t	1.863	1.863
df	8	7.724
Sig. (2-tailed)	.099	.101
Mean Difference	218.60000	218.60000
Std. Error Difference	117.31709	117.31709
95% Confidence Interval of the Difference		
- Lower	-51.93370	-53.62527
- Upper	483.13370	490.82527

The equality of variances between B40 and M40 data should be considered by determining the significance value (p-value). If the p-value is above 0.05, the data assume equal variances. The p-value of the data 0.815, which are higher than 0.05, therefore equal variances are considered. When equal variance is considered, the p-value for t-test of Equality of Mean is being discussed. The data assumed as equal variance; therefore, only Sig. (2-tailed) is studied. The p-value of the Sig.2 (tailed) is 0.099, which is greater than 0.05. Therefore the alternative hypothesis is rejected, and the null hypothesis is accepted. Therefore, there is no difference between B40 and M40 on positive sentiments tweets.

Inspection of the descriptive study found out that homogeneity of variance as assessed by Levene's Test for Equality of Variances. Therefore, an independent t-test was run on the data, with a 95% confidence interval (CI) for the mean difference. It was found that after the two interventions, the polarity of positive sentiments tweets in B40 was not significantly different than polarity of positive sentiments in M40 ($t(8) = 1.863$, $p = 0.099$).

Table 3. T-test of social classes group and positive sentiments

	neg_tweets	
Levene's Test for Equality of Variances	**Equal variances assumed**	**Equal variances not assumed**
F	1.348	
Sig.	.279	
t-test for Equality of Means		
t	3.660	3.660
df	8	6.628
Sig. (2-tailed)	.006	.009
Mean Difference	210.40000	210.40000
Std. Error Difference	57.48478	57.48478
95% Confidence Interval of the Difference		
- Lower	77.83986	72.90704
- Upper	342.96014	347.89296

The equality of variances between B40 and M40 data should be considered by determining the significance value (p-value). If the p-value is

above 0.05, the data assume equal variances. The p-value of the data 0.279, which are higher than 0.05, therefore equal variances are considered. When equal variance is considered, the p-value for t-test of Equality of Mean is being discussed. The data assumed as equal variance; therefore, only Sig. (2-tailed) is studied. The p-value of the Sig.2 (tailed) is 0.006, which is lower than 0.05. Therefore the null hypothesis is rejected, and the alternative hypothesis is accepted.

4.2 Display the Results in Dashboard

Fig. 7. The dashboard of social classes group B40 and numbers of tweets sentiments

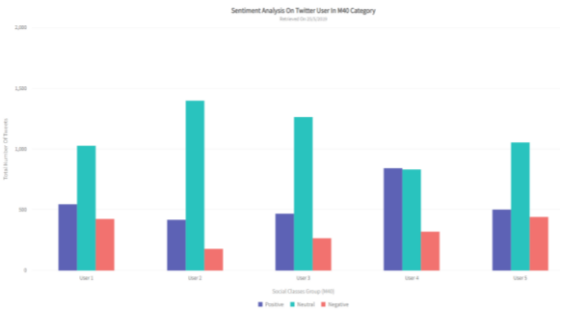


Fig. 8. The dashboard of social classes group M40 and numbers of tweets sentiments

4.4 Future Works

There a few suggestions for the improvement of research studies arise. Firstly, the dataset collected could be more precise in normal distributions. Next, the machine learning approach should be added to precisely classify tweet into positive and negative sentiments based on a pre-determined training dataset. Also, the approach should accept more than one language of words corpus to add more lexicon-based dictionary. Lastly, the dataset should filter more noise, such as abbreviations and slangs, along with stemming and lemmatizing words. For future works, the process of learning based on polarity sentiments should be improved. Machine learning would be added to train all the dataset with suitable polarity values. Besides, the tweets will be classified by Natural Language Processing to avoid ambiguity and neutral sentiments of classification

References

[1] Diakopoulos, N., Naaman, M., & Kivran-Swaine, F. (2010). Diamonds in the rough: Social media visual analytics for journalistic inquiry. VAST 10 - IEEE Conference on Visual Analytics Science and Technology 2010, Proceedings, 115–122. http://doi.org/10.1109/VAST.2010.5652922.

[2] Semega, J. L., Fontenot, K. R., & Kollar, M. A. (2017). Income and Poverty in the United States: 2016. US Census Bureau, (September), P60-249. http://doi.org/Retrieved from http://www.census.gov/content/dam/Census/library/publications/2016/demo/p60-256.pdf

[3] Khazanah Research Institute. (2014). The State of Households, 1–68.

[4] REPORT OF HOUSEHOLD INCOME AND BASIC AMENITIES SURVEY 2016. (2017, October 9). Retrieved February 25, 2019, from https://www.dosm.gov.my/v1/index.php?r=column/pdfPrev&id=RUZ5REwveU1ra1hGL21JWVlPRmU2Zz09.

[5] Cao, Y., Zhang, P., & Xiong, A. (2015). Sentiment Analysis Based on Expanded Aspect and Polarity-Ambiguous Word Lexicon. IJACSA) International Journal of Advanced Computer Science and Applications, Citeseer, 6(2), 97–103

[6] Singh, T., & Kumari, M. (2016). Role of Text Pre-processing in Twitter Sentiment Analysis. Procedia Computer Science, 89, 549-554. doi:10.1016/j.procs.2016.06.095

[7] [22] Oliveira, J. E., Cotacallapa, M., Seron, W., Santos, R. D., & Quiles, M. G. (2016). Sentiment and Behavior Analysis of One Controversial American Individual on Twitter. Neural Information Processing Lecture Notes in Computer Science, 509-518.

doi:10.1007/978-3-319-46672-9_57

[8] Shalev-Shwartz, S., & Ben-David, S. (2014). Understanding Machine Learning: From Theory to Algorithms. http://doi.org/10.1017/CBO9781107298019

[9] Diakopoulos, N., Naaman, M., & Kivran-Swaine, F. (2010). Diamonds in the rough: Social media visual analytics for journalistic inquiry. VAST 10 - IEEE Conference on Visual Analytics Science and Technology 2010, Proceedings, 115–122. http://doi.org/10.1109/VAST.2010.5652922

[10] Njoroge, R. I. T. A. (2013). Impacts of social media among the youth on behavior change: a case study of University students in selected universities in Nairobi, Kenya. , 80.

[11] Phua, Y. (2013). Social media sentiment analysis and topic detection for Singapore English, (September). Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA595207

[12] Saif, H., Fernandez, M., He, Y., & Alani, H. (2013). Evaluation datasets for Twitter sentiment analysis a survey and a new dataset, the STS-Gold. CEUR Workshop Proceedings, 1096, 9–21

[13] Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in Facebook and its application to e-learning. Computers in Human Behavior, 31(1), 527–541. http://doi.org/10.1016/j.chb.2013.05.024

[14] Vural, A. G., Cambazoglu, B. B., & Karagoz, P. (2014). Sentiment-focused web crawling. ACM Transactions on the Web (TWEB), 8(4), 22

[15] Zamani, N. A. M., Abidin, S. Z. Z., Omar, N., & Abiden, M. Z. Z. (2014). Sentiment Analysis : Determining People ' s Emotions in Facebook 2 Related Work. Applied Computational Science, ISBN: 978-, 111–116

[16] Ahmad, N. F., Mansor, M., & Paim, L. (2016). Income poverty and well-being among vulnerable households: A study in Malaysia. Asian Social Science, 12(2), 195– 204. http://doi.org/10.5539/ass.v12n2p195

[17] Christofferson, J. P. (2016). How is Social Networking Sites Effecting Teen ' s Social and Emotional Development : A Systemic Review.

[18] Havener, L. S., & Allen, N. B. (2016). THE EFFECTS OF SOCIAL MEDIA AND SOCIAL NETWORKING SITE USAGE ON THE MENTAL HEALTH AND WELLBEING OF ADOLESCENTS Title: The Effects of Social Media and Social Networking Site Usage on the Mental Health and Wellbeing of Adolescents Background: Many questions, (September).

[19] Muhammad A: Contextual Lexicon-based Sentiment Analysis for Social Media. (2016).

[20] Dani, H., Li, J., & Liu, H. (2017, September). Sentiment Informed Cyberbullying Detection in Social Media. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases (pp. 52-67). Springer, Cham

[21] Semega, J. L., Fontenot, K. R., & Kollar, M. A. (2017). Income and Poverty in the United States: 2016. US Census Bureau, (September), P60-249. http://doi.org/Retrieved from http://www.census.gov/content/dam/Census/library/publications/2016/demo/p60-256.pdf

[22] Sarlan, A., Nadam, C., & Basri, S. (2014). Twitter sentiment analysis. Proceedings of the 6th International Conference on Information Technology and Multimedia. doi:10.1109/icimu.2014.7066632

[23] Shen, C. (n.d.). Text Analytics of Social Media: Sentiment Analysis, Event Detection and Summarization. doi:10.25148/etd.fi14110776

A Novel Approach for Airport Recommendation Using the Comparison of Hierarchical Regression, PLS-SEM, and Bayesian Network-Based PSEM

Cheong Kim[1,2], Dong Eun Lee[2], Kyun Sun Eo[2], Kun Chang Lee[2,3*]
[1]Manager, Predictive Analytics and Data Science, ACI World
Montréal, Québec, H4Z 1G8, Canada
[2] SKK Business School, Sungkyunkwan University
Seoul 03063, Republic of Korea
[3]Department of Health Sciences & Technology,
SAIHST (Samsung Advanced Institute for Health Sciences & Technology)
Sungkyunkwan University
Seoul 03063, Republic of Korea
[e-mail: saga@g.skku.edu, leede93@gmail.com, eokyunsun@gmail.com, kunchanglee@gmail.com]
*Corresponding author: Kun Chang Lee

Abstract

The overarching philosophy of conducting this research is, first, to reveal the most significant factors that affect consumers' recommendations upon airports, second, to find out if three different methodological approaches, such as conventional hierarchical regression and SEM versus Bayesian network-based PSEM, differ the results. A total of 11,656 consumer reviews for 920 airports worldwide were acquired from an airport-review website. The dataset was analyzed using hierarchical regression, PLS-SEM as well as unsupervised Bayesian network-based PSEM to validate the hypotheses. All of the three methodologies confirmed that consumers' perception of the quality of servicescape, service, and airport staff factors significantly influences the subjects' intention to recommend the airports. However, only the PSEM results showed that wifi connectivity doesn't impact recommendation, and it formulized a distinct model structure compared to PLS-SEM. This research provides practical implications to airport operators that enhancing three quality dimensions will induce consumers to have positive intentions to recommend their airports and thus have a higher possibility of retaining more passengers. Considering that the power of e-WOM (word of mouth), practitioners in the airport management field could take these online review sites to mitigate consumers' dissatisfaction and improve facilitation with the guideline that this research suggests. This study revealed the impact of crucial factors that affects airport recommendation by comparing three different methodologies; hence, it provides a more rigorous implication that could be employed both in academia and the field.

***Keywords*:** Airport reviews, Service quality, Recommendation, Hierarchical regression, SEM, Bayesian network

1. Introduction

In most industries, consumer satisfaction and firms' revenue are crucially related. In seeking to find an adequate approach to convene consumers' satisfaction, firms have invested endless budget and effort. Aviation is one of the most sensitive industries affected by consumer satisfaction, considering that interaction with human agents, service, and servicescape is fundamental during the entire consumer experience of using the service. Therefore, consumers' positive or negative perceptions of service experience are exceedingly vital regardless of the aviation market's stakeholders.

Airports are not an exception. Air travelers have consequential preferences among airports, and there is ever-increasing importance for the field practitioners to segregate their airports by satisfying consumer insights in order to acquire competitive advantages in the market [1]. One of the frequently utilized media channels in this activity for airport operators' survival is a website review [2]. Because of the powerful influence of e-WOM (word of mouth) for air travelers [3], several airports in the world use consumers' reviews on online sites as one of the fundamental references to enhance their service quality and be recommended. Therefore, it would be fruitful to discover what factors most affect airport users' satisfaction by using relevant online reviews and appropriate approaches.

Several researchers tried to prove this topic; most of them used a conventional regression or structural equation model [1, 3, 4]. However, for airport field practitioners, these conventional methodologies are complicated to interpret and challenging to provide practical implications that could be directly applied to their works. In this research, we have employed and compared three different approaches for analyzing consumers' online reviews that affect airport recommendation: hierarchical regression, partial least square structural equation model (PLS-SEM), and Bayesian network-based probabilistic structural equation model (PSEM). By doing so, this research provided both academic and practical implications in the field of airport service quality. The rest of this paper is organized as follows: theoretical background, study 1, 2, and 3, and conclusion.

2. Theoretical background and hypotheses development

2.1 Servicescape and service

Servicescape (i.e., check-in and security facilities, terminal signs, terminal cleanliness, and terminal seating) include the physical and environmental components that the airport operators manage to facilitate users' experience at airports. Previous research has suggested the importance of the physical and environmental criteria on users' perceptions of service quality [5-7].

Next, *service* (i.e., food and beverage, airport shopping, and wifi connectivity) is what airport users can spend time with while they wait for their flights [1]. Since every airport needs users' physical presence and a time commitment, these requirements in terms of servicescape and service could significantly impact users' quality perceptions [1]. These factors would also be the fundamental elements that the airport users perceived benefits from their experience and thus induce them to use the service repeatedly [8]. Based on this theoretical backgrounds, the below hypothesis is proposed:

H1. Servicescape will positively influence airport recommendation of users in online reviews.

H2. Service will positively influence airport recommendation of users in online reviews.

2.2 Airport staff

The last influence on airport users' intention of recommendation is interaction with the *airport staff*. Considering most user activities at airports are closely related to the service personnel (i.e., airline agent, security agent, information desk clerk), their attitude, kindness, and responsiveness would be a direct indicator for users to perceive airport service quality. Researchers also articulated the importance of service providers' interaction with consumers [1, 9, 10]. Thus, the third hypothesis in this research is proposed:

H3. Airport staff will positively influence airport recommendation of users in online reviews.

3. Study 1

In this study, we obtained consumer review scores and recommended an airport. The dataset was crawled and obtained on airlinequality.com that provides customer reviews on worldwide airports. The data sample contained 11,656 consumer review ratings with the following attributes: *terminal signs* (*TSI*), *terminal cleanliness* (*TCL*), *terminal seating* (*TSE*), *food and beverages* (*FNB*), *airport shopping* (*ASH*), *wifi connectivity* (*WCO*), *airport staff*, and *recommendation*.

We conducted a hierarchical regression analysis using SPSS 22 to discover the independent variables' significance to the dependent variable. First, we inserted *servicescape* groups, including *terminal signs*, *terminal cleanliness*, and *terminal seating*. Then, we included the quantitative ratings for *service* group (*food and beverages*, *airport shopping*, *wifi connectivity*) and *airport staff* in the model..

In first set of models (Model 1.1), the ratings on *terminal seating* (β = .113, p < 0.001) has the most significant relationships with the dependent variable than the other variables. In the next set of models (model 1.2) with *service* group, *terminal seating* still showed the most substantial impact on *recommendation* (β = .081, p < 0.001). Finally, in the last model with *airport staff* (model 1.3), *terminal seating* was changed to the second most impactful variable (β = .063, p < 0.001), while *airport staff* became the first (β = .116, p < 0.001). As we involve additional groups, the adjusted R^2 improved from 0.532 (model 1.1) to 0.619 (model 1.3), suggesting that the models' variance explanation has been enhanced. In addition to the coefficient tests, we evaluated the multicollinearity of each of the models. The results suggested that all the variance inflation factors are lower than 4.0, as **Table 1** indicates; thereby, multicollinearity was not found in the model.

Table 1. The results of hierarchical regression

DV: Recommendation		Model 1.1	Model 1.2	Model 1.3	VIF
	Adjusted R^2	.532***	.556***	.619***	
	(Constant)	-.394***	-.439***	-.426***	
	Terminal signs	.093***	.071***	.043***	2.463
Servicescape	Terminal cleanliness	.054***	.034***	.012***	2.846
	Terminal seating	.113***	.081***	.063***	3.041
	Food & beverages		.042***	.031***	3.441
Service	Airport shopping		.020***	.011**	3.117
	Wifi connectivity		.036***	.017***	1.843
Airport staff				.116***	2.267

*p < .05, **p < .01, ***p < .001

4. Study 2

In Study 2, we use the identical dataset as in study1. To analyze the research model, a structural equation model (SEM) with SmartPLS 3.0 was used. SEM has been extensively applied to verify hypothesis testing results, including consumer satisfaction and service quality-related studies [11-13]. Partial least square (PLS) is a suitable methodological approach to discover relationships between constructs based constructs built on a hypothesis [14].

We applied two formative second-order constructs: servicescape (with three dimensions: *terminal signs*, *terminal cleanliness*, *terminal seating*) and service (with three dimensions: *food & beverages*, *airport shopping*, *wifi connectivity*). The repeated indicator approach was employed for analyzing these formative second-order constructs. With this approach, we can appropriately predict the second-order formative constructs without any flooding-out influence by implementing this two-stage approach [15].

In order to evaluate the significance of the second-order formative constructs, bootstraping with 1,000 samples was employed to assess the upward dimensional effects.

The results suggest that every first-order construct (*TSI*, *TCL*, *TSE*, *FNB*, *ASH*, *WCO*, and *staff*) that was incorporated in the second-order formative constructs (*servicescape* and *service*) are significantly affecting, as presented in Table 2. Also, the heterotrait–monotrait (HTMT) ratio was utilized to evaluate discriminant validity. Table 3 demonstrated that the results of the HTMT ratio evaluation were smaller than 1.00. For discriminant validity, this is accepted and recommended criteria [16].

Table 2. Discriminant validity using the heterotrait–monotrait (HTMT) ratio.

	Original β	*Mean β*	*t-value*	*p-value*
TCL -> Servicescape	.371	.371	343.355	.000
TSE -> Servicescape	.380	.379	341.574	.000
TSI -> Servicescape	.366	.366	355.781	.000
FNB -> Service	.393	.393	253.597	.000
ASH -> Service	.350	.350	248.842	.000
WCO -> Service	.401	.401	260.728	.000
Service -> Recomm	.161	.161	13.980	.000
Servicescape -> Recomm	.326	.326	25.920	.000
Staff -> Recomm	.375	.375	35.173	.000

The test result of hypothesis 1 found that servicescape significantly positively influence *recommendation* ($\beta = 0.326$, $p < 0.000$). In addition to the test result of hypothesis 1, the test result of hypothesis 2 also indicated that the *service* is significantly positively related to the *recommendation* ($\beta = 0.161$, $p < 0.000$). Finally, the test results of hypothesis 3 indicated that *airport staff* significantly positively impacts *recommendation* ($\beta = 0.375$, $p < 0.000$). Hence, all three hypotheses in this research were supported, as presented in **Fig. 1**.

Table 3. Discriminant validity using the heterotrait–monotrait (HTMT) ratio.

Construct		TCL	TSE	TSI	ASH	FNB	WCO	Staff	Recommendation
Servicescape	TCL								
	TSE	.746							
	TSI	.683	.677						
Service	ASH	.621	.653	.618					
	FNB	.641	.699	.628	.804				
	WCO	.590	.586	.561	.549	.562			
Airport staff		.650	.648	.654	.590	.608	.577		
Recommendation		.630	.674	.646	.591	.621	.552	.721	

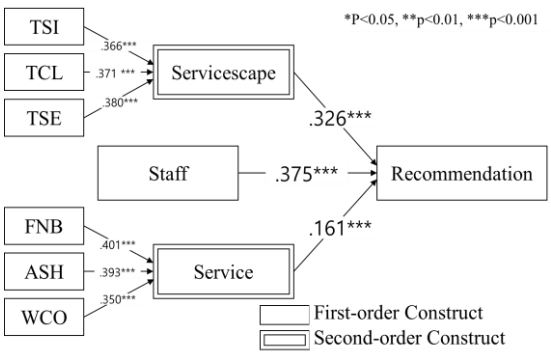


Fig. 1. Results of the hypotheses tests

5. Study 3

The probabilistic structural equation model (PSEM) is similar to traditional SEM [17]. However, PSEM is based on a Bayesian network structure a contrast to a series of equations. PSEM has three fundamental characteristics distinguished from SEM. First, all relationships in a PSEM are probabilistic. Second, PSEM is nonparametric, which facilitates the representation of nonlinear relationships between categorical variables. Last, the structure of PSEM is partially or fully machine-learned using the dataset.

We attempted five different algorithms (i.e., Taboo [18], Sopleq [19], EQ [20], Maximum Spanning Tree [21], and TabooEQ [17] with the dataset to learn a network using Bayesia 9 to find the best algorithm with the smallest MDL scores output. Minimum description length (MDL) scores measure to optimize while searching for the best possible network. Minimizing this score comprises finding the most appropriate algorithmic approach [22]. Therefore, the MDL score has to be minimized to obtain the best solution. The result of MDL indicated that TabooEQ, an unsupervised structural learning algorithm based on EQ [17] is the most appropriate algorithm for the model; the initial MDL score was 177,329.315 and became the lowest with the score of 162,116.491 after applying Taboo EQ.

Next, we conducted variable clustering, and four groups, *servicescape*, *service*, *wifi connectivity*, and *guidance*, were formulated. In this step, we find a difference between the PLS-

SEM and the PSEM. The two second-order constructs in PLS-SEM has three first-order descendant constructs; on the other hand, PSEM built three groups with two sub-variables and one individual group. Next, we conducted multiple clustering as well as the second unsupervised learning with Taboo EQ algorithm to finalize the model, as shown in **Fig. 2**. *Wifi connectivity* factor did not have any relationship with the target node. This is another distinguishable finding in PSEM compare to PLS-SEM.

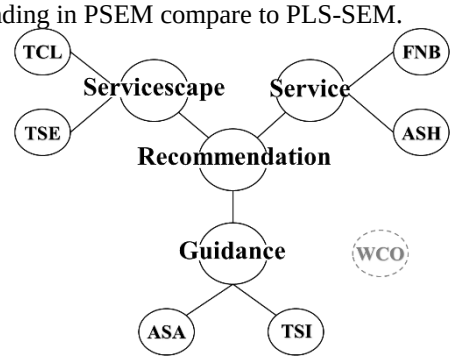


Fig. 2. Bayesian network-based PSEM

The PSEM standardized total effect (STE) results immediately highlight the order of importance of the factors relative to the target node (*recommendation*). *Guidance* is the most crucial factor affecting the target node, with the STE score of 0.72. *Servicescape* and *service* factors are following after *guidance*.

In addition to STE analysis, we analyze the probabilistic causes with genetic optimization that are strong arbitrary predictors of the target node. The results discovered that the best solution would be receiving the highest scores for all three indicators (*servicescape*, *service*, *guidance*) that will finally drive *recommendation* with the probability of 92.50%.

6. Conclusion

This paper investigated the relationships between *servicescape*, *service*, *airport staff*, and consumers' intention toward airports' *recommendations* in detail, using three different sorts of methodological approaches: hierarchical regression, PLS-SEM, and PSEM.

The significant contribution of this research would be recapped in three pillars. First, we revealed that *wifi connectivity* is an insignificant variable only in the PSEM model. This difference is because the wifi at airports is more likely considered a compulsory facility considering that smartphones are now an everyday device for consumers. Therefore, no wifi might cause critical damage to airport users' quality perception.

Also, in the PSEM model, the variable terminal signs were grouped with *airport staff* as *guidance* factor, unlikely to what we have expected in the hypotheses development and the results in PLS-SEM. It is because airport users contemplate *airport staff* and *terminal signs* for the same purpose: *airport staff* and *terminal signs* provide relevant information about the airports as a guidance apparatus. This difference in the results is because of the Bayesian network's distinct algorithm that is purely performed with probabilistic statistics and machine learning. It would be hasty to conclude that the result from the PSEM is superior to the results of hierarchical regression or PLS-SEM since human intervention with the experts' domain knowledge would also be quite crucial for analyzing the phenomenon. However, PSEM would be an excellent supplementary instrument for researchers and field practitioners to confirm the other methodologies' test results.

Lastly, Bayesian network-based PSEM can provide appropriate resolutions to solve the problematic issue, such as how to increase consumer satisfaction, service quality, and recommendation, with the genetic approach that is relatively easy to interpret for managers. Especially in this research, receiving the highest scores in all three criteria, *servicescape*, *service*, *guidance* from airport users will induce them to make *recommendations* for the airports with 92.50% of probability.

References

[1] D. Fodness, and B. Murray, "Passengers' expectations of airport service quality," *Journal of Services Marketing*, pp., 2007.

[2] L. Martin-Domingo, J.C. Martín, and G. Mandsberg, "Social media as a resource for sentiment analysis of Airport Service Quality (ASQ)," *Journal of Air Transport Management*, Vol 78, pp. 106-115, 2019.

[3] R. Filieri, and F. McLeay, "E-WOM and accommodation: An analysis of the factors that influence travelers' adoption of information from online reviews," *Journal of Travel Research*, Vol 53. No.1, pp. 44-57,

2014.

[4] S.-J. Hong, D. Choi, and J. Chae, "Exploring different airport users' service quality satisfaction between service providers and air travelers," *Journal of Retailing and Consumer Services*, Vol 52, pp. 101917, 2020.

[5] M.E. Rys, J.O. Fredericks, and D. Luery, "Value= quality? Are service value and service quality synonymous: A decompositional approach," *Add Value to Your Service*, pp. 25-28, 1987.

[6] P.A. Dabholkar, D.I. Thorpe, and J.O. Rentz, "A measure of service quality for retail stores: scale development and validation," *Journal of the Academy of Marketing Science*, Vol 24. No.1, pp. 3, 1996.

[7] M.K. Brady, and J.J. Cronin Jr, "Some new thoughts on conceptualizing perceived service quality: a hierarchical approach," *Journal of Marketing*, Vol 65. No.3, pp. 34-49, 2001.

[8] C. Kim, K.C. Lee, and F.J. Costello, "The Intention of Passengers Towards Repeat Use of Biometric Security for Sustainable Airport Management," *Sustainability*, Vol 12. No.11, pp. 4528, 2020.

[9] J.J. Cronin Jr, and S.A. Taylor, "Measuring service quality: a reexamination and extension," *Journal of Marketing*, Vol 56. No.3, pp. 55-68, 1992.

[10] H. Oh, and S.C. Parks, "Customer satisfaction and service quality: a critical review of the literature and research implications for the hospitality industry," *Hospitality Research Journal*, Vol 20. No.3, pp. 35-64, 1996.

[11] A. Geebren, A. Jabbar, and M. Luo, "Examining the role of consumer satisfaction within mobile eco-systems: Evidence from mobile banking services," *Computers in Human Behavior*, pp. 106584, 2020.

[12] H.-J. Chen, "Linking role definition rigidity to elementary school teachers'e-service for in-service teacher development," *Computers in Human Behavior*, Vol 107, pp. 105990, 2020.

[13] N. Subramanian, A. Gunasekaran, J. Yu, J. Cheng, and K. Ning, "Customer satisfaction and competitiveness in the Chinese E-retailing: Structural equation modeling (SEM) approach to identify the role of quality factors," *Expert Systems with Applications*, Vol 41. No.1, pp. 69-80, 2014.

[14] C.R. Fornell, and D.F. Lacker, "Two structural equation models with unobservable variables and measurement error," *Journal of Marketing Research*, Vol 18. No.1, pp. 39-50, 1981.

[15] W.W. Chin, B.L. Marcolin, and P.R., Newsted, "A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study," *Information Systems Research*, Vol 14. No.2, pp. 189-217, 2003.

[16] J. Henseler, C.M. Ringle, and M. Sarstedt, "A new criterion for assessing discriminant validity in variance-based structural equation modeling," *Journal of the Academy of Marketing Science*, Vol 43. No.1, pp. 115-135, 2015.

[17] S. Conrady, and L. Jouffe, "Introduction to bayesian networks & bayesialab," *Bayesia SAS*, pp., 2013.

[18] D. Chickering, D. Geiger, and D. Heckerman, "Learning Bayesian networks: Search methods and experimental results," *in Proc. of Fifth Conference on Artificial Intelligence and Statistics*, pp. 112-128, 1995.

[19] M. Zargoush, F. Alemi, V.E. Vinzi, J. Vang, and R. Kheirbek, "A psychological approach to learning causal networks," *Health care Management Science*, Vol 17. No.2, pp. 194-201, 2014.

[20] P. Munteanu, and Bendou, M., "The EQ framework for learning equivalence classes of Bayesian networks," *in Proc. 2001 IEEE International Conference on Data Mining*, pp. 417-424, 2001.

[21] W. Lam, and F. Bacchus, "Learning Bayesian belief networks: An approach based on the MDL principle," *Computational Intelligence*, Vol 10. No.3, pp. 269-293, 1994.

[22] J.R. Lechien, C.M. Chiesa - Estomba, S. Place, Y. Van Laethem, P. Cabaraux, Q. Mat, K. Huet, J. Plzak, M. Horoi, S. Hans, and M.R. Barillari, "Clinical and epidemiological characteristics of 1,420 European patients with mild- to- moderate coronavirus disease 2019," *Journal of Internal Medicine*, pp., 2020.

C3BR: A Novel Multi-Label Classification Ensemble Mechanism

Jong Yoon Won[1], Chang Min Kang[1], Dong Eun Lee[1], Kun Chang Lee[1,2*]
[1]SKK Business School, Sungkyunkwan University
Seoul, South Korea
[2]Professor, Samsung Advanced Institute for Health Sciences & Technology (SAIHST)
Sungkyunkwan University
Seoul 03063, South Korea
[e-mail: yoonjbest1@gmail.com, 77aktrp@gmail.com, leede93@gmail.com, kunchanglee@gmail.com]
*Corresponding author

Abstract

This paper aims to propose a novel multi-label classification mechanism entitled C3BR. The proposed C3BR stands for 'ensemble Classifier Chain applied, Correlation feature selection based on the Binary Relevance approach', and it is applied to a big dataset from airline service review to verify its validity. The main contributions of the proposed ensemble C3BR are threefold. Firstly, different from conventional multi-label classification (MLC) methods, the C3BR adopts feature selection (FS) mechanism to reduce the number of features to be considered during classification procedures. Therefore, the C3BR helps to relieve computational burden. Secondly, our proposed C3BR ensembles an MLC method called CC(Chain Classifier) with those features selected from integrating correlation feature selection method with BR (Binary Relevance) MLC. Thirdly, the proposed C3BR is compared with other available benchmarking MLC methods like BR (binary relevance), CC (chain classifier), FW(four-class pairwise), RAkEL(Random k-Label sets), MCC(monte carlo classifier chain), CDN(conditional dependency network), LP (label powersets). Based on the airline service review dataset, we proved the empirical validity of the proposed C3BR.

***Keywords*:** Multi-Label Classification, Feature Selection, Binary Relevance, Correlation-based Feature Selection

1. Introduction

Traditional machine learning method aims to find the most relevant label (target variable). It is called single-label learning algorithm. Unlike the single-label learning, multi-label learning algorithm predicts the set of relevant label(two or more target variables: Y). In other words, multi-label classification is a special learning problem, in which one instance can possibly have several label-set at the same time [1]. Recently, in the field of artificial intelligence and machine learning for real applications, multi-label classification has been increasing interest among scholars [2]. Also, it is intergraded in various real-world applications, such as Text categorization [3], Behavior recognition [4], Gene annotation [5], and more. Researches are being conducted in collaboration with the humanities department, using human information [6]. This shows that multi-label classification applies extensively to other real-world problems. As multi-label data set includes a vast number of features, it might be irrelevant, redundant, and can lead to deteriorated

classification performance which increases the computational cost. [7]. To address this problem, we apply the feature selection method.

Feature selection is a technique that seeks to select meaningful features in order to eliminate irrelevant features and to maintain as much relevent information as possible. [2]. As multi-label learning has attracted more attention, previous studies stressed the relevance of the feature selection problem in multi-label learning. Even though a multi-label feature selection task is worthwhile than a single-label feature selection task, it is considered to be more complicated and hard [8]. To solve this problem, a commonly encountered method is to transform the multi-label problems into numerous single-label problems. Then the relevant features for each converted new single label can be clearly selected with the benefit of any traditional single-label methods [9].

As an essential preprocessing tool, feature selection is a crucial step in machine learning. It can enhance classification accuracy and effectively reduce the dataset dimensionality by removing unnecessary features. A huge amount of research previously concentrated on feature selection for algorithm adaptation. However, only few studies discussed the problem of multi-label data in problem transformation. Therefore, we propose a C3BR method for multi-label feature selection based on problem transformation.

2. Previous Studies

2.1 Multi-Label Classification

The multi-label classification (MLC) refers to a special type of classification in which the dataset is composed of multiple labels. For example, the single-label classification(called machine learning) usually deals with a dataset in which inputs dataset is always related to single label, i.e., single dependent variable y. Therefore, the SLC problem is denoted as:

$$\begin{pmatrix} x_{1,1} & \cdots & x_{1,m} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,m} \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

In contrast, MLC problem is denoted as:

$$\begin{pmatrix} x_{1,1} & \cdots & x_{1,m} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,m} \end{pmatrix} = \begin{pmatrix} y_{1,1} & \cdots & y_{1,l} \\ \vdots & \ddots & \vdots \\ y_{n,1} & \cdots & y_{n,l} \end{pmatrix}$$

Compared with the SLC, the MLC has been known to be hard for decision-makers to utilize in their real situations. Its reasons are as follows. Firstly, since it has multiple labels, the MLC training procedures are separated into two methods such as Problem Transformation methods, and Algorithm Adaptation methods [10].

The former converts the MLC problem into a number of SLC problems [11]. The latter suggests a new algorithm to handle the multi-label dataset effectively [12]. However, the Algorithm Adapatiom methods propose complicated computational procedures fitting into the characteristics of the MLC problems, that causes serious troubles to most decision-makers working in the management fields. Therefore, this study prefers the Problem Transformation methods which converts original MLC problems into a series of SLC problems and allows conventional training methods to be more easily applied. For example, in the context of MLC problems, many training methods are available like BR (binary relevance), CC (chain classifier), FW(four-class pairwise), LP (label powersets), MCC(Monte Carlo classifier chain), CDN(conditional dependency network), RAkEL(Random k-Labelsets) and Ensemble-CC methods.

2.2 Multi-Label Feature Selection

Feature selection is a data preprocessing step in multi-label classification, it reduces the feature space dimension, and enhances the algorithm performance [13]. It removes unnecessary features and only keeps the features that have sturdy category discrimination ability to construct the feature set. Multi-label feature selection procedures can be separated into three categories: wrapper, filter, and embedded method [14]. Compared to other techniques, filter methods are simple. These methods have lower computational complexity, which can make results more suitable [15]. Besides, they require a statistical analysis of a feature set without applying any learning algorithm. Therefore, this study focuses on the filter feature selection strategy. We applied Correlation Feature Selection(CFS), ReliefF(RF), Information

Gain(IG), and CV(Cross Validation) to perform multi-label feature selection of the problem transformation method approach.

Feature selection is a drastic research topic in supervised, and non-supervised machine learning. Feature dimensionality reduction algorithms can be separated into two concepts; feature extraction, and feature selection.
Nevertheless, almost all of the studies about feature selection support single-label classification, and few results are related to multi-label classification using BR, LP, and Pruned Problem Transformation (PPT), which is also known as the problem transformation method [16]. When using previous algorithms, the results can be solved simply by a single-label learning algorithm. For instance, BR transforms each label into a single binary label form, and then feature subgroup selected with RF and IG measures. However, this method may yield poor classification performance because it is handling the labels independently, and causing information losses, in case of correlation between labels [17]. To sum up, the problem transformation method does not consider the label correlation into account, which could lower the classification performance. Therefore, some feature selection methods are advancing to select features on a multi-label dataset straightly. The similarity between features and labels is calculated based on the Laplacian score. Zhang et al. offered a novel multi-label feature selection algorithm that tries to use manifold regularization to find the features with maximum predictive power [18]. Information-theoretic feature ranking algorithm, which is one of the multi-label feature selection algorithms was proposed, where the top-ranked features were selected from the list in which all features ranked by their importance. Almost all methods select features directly from a multi-label dataset by considering the correlation between labels. The selected features have specific, meaningful information for each label; however, do not have any distinct value for another label which has a little dependency on the label. Meanwhile, it is difficult to analyze the dependency between each pair of labels when datasets have a considerable number of labels.

3. Proposed Method

3.1 C3BR Model

The proposed C3BR model in this study is as follows. C3BR aims to analyze most effectively for multi-label learning. Therefore, we propose a multi-label ensemble mechanism that combines feature selection and classification ensemble algorithms. The C3BR combination's step is composed of below.
First, we suggest a feature selection method by merging the BR method with the CFS method, which is a problem transformation method of multi-label learning. Second, the performance is expected to be improved when applying Ensemble CC. By combining the two steps from above, we recommend an effective multi-label feature selection analysis mechanism.

3.2 Data collection and procedure

We collected the data from the airlinequality.com website, which includes customer reviews on various airlines worldwide. After removing missing data, the total sample data contains 18,498 airline customer reviews on 512 different airlines—the customer reviews posted between 2005 and 2019. The dataset includes the total number of airline recommendations, textual reviews, and aspect-wise ratings of the following categories: cabin staff behavior, seat comfort, food and beverages, ground service, and value for money, etc. We conducted multi-label learning based on the data using the following steps. First, eight target values were set by designating aspect-wise ratings as target variables. Second, an instance was constructed by embedding the online review.

4. Result

Hamming Loss is a commonly used indicator for measuring Multi-Label Classification performance. In this study, Hamming Loss was used as a major indicator, and accuracy was used as an auxiliary indicator. The results of the study are shown in Table 1. The C3BR model, which is presented in this study is highlighted. Next to the result, numbers written in parenthesis are the representation of the order of excellence of this corresponding algorithm.
Next, performance will be compared before and after feature selection. Looking at the accuracy of

the baseline, the accuracy after the feature selection generally showed high performance. Similarly, Hamming Loss also showed excellent overall results after feature selection. This demonstrates that feature selection does not degrade the performance of the data, even though the number of attributes was reduced. Second, the CFS method was shown to be the most effective feature selection method. **Table 2** shows the mean of the results. The CFS in the BR approach showed the highest accuracy (0.652) and lowest Hamming Loss (0.187). Third, if we compare Multi-Label Classification algorithms, Ensemble CC has shown the best classification performance. By looking at the results comprehensively, the C3BR presented in this study proved to be a mechanism for effective Multi-Label learning.

Table 1. Mean of Table 2

		Accuracy	Hamming Loss
BR Base	**CFS**	.652	.187
	IG	.605	.208
	ReliefF	.602	.210
	CV	.578	.232
CC Base	**CFS**	.522	.274
	IG	.605	.208
	ReliefF	.598	.210
	CV	.583	.228
LP Base	**CFS**	.652	.194
	IG	.649	.195
	ReliefF	.605	.205
	CV	.597	.215

5. Conclusions

This paper proposes and analyses novel mechanisms for multi-label learning, which uses the filter approach of feature selection methods and ensemble CC approach. At the results, two specification multi-label feature selection approaches were used.

First, a multi-label dataset was transformed into a single-label dataset by using BR problem transformation approach. Next, we performed feature selection by the CFS method based on Best-First Search what after a new multi-label data set was constructed with its subset. Second, we applied the ensemble CC, which improves multi-label learning efficiency. To measure the goodness of features, we used eight multi-label algorithms and three feature selection algorithm benchmarks.

The results showed that the methods in which CFS used as a feature evaluation measure, often select a smaller number of features than RF, IG, and CV with no degradation on the correspondent classifiers. That could be since CFS considers the relationships between features. Finally, we completed the multi-label learning analysis mechanism by applying the ensemble CC. The ensemble CC classifier proved to be superior to other algorithms; hence, the suggested multi-label feature selection method is useful for the future compared to FS methods to side with multi-label learning. Besides, the mechanism applied to the classification algorithm increased the effectiveness of the analysis. In future work, we make a plan that expands the experimental evaluation of deep-learning using synthetic datasets. Furthermore, to solve the multi-label feature selection problem directly, considering the analysis of potential CFS extensions without any problem transformation, will also be researched.

Acknowledgment: This work was supported by the Technology development Program (S2926521) funded by the Ministry of SMEs and Startups (MSS, Korea).

References

[1] M. L. Zhang, Z. H. Zhou, "A review on multi-label learning algorithms." *IEEE transactions on knowledge and data engineering*, Vol 26, No. 8, pp. 1819-1837 2013.

[2] L. Sun, T. Yin, W. Ding, Y. Qian, J. Xu, "Multilabel Feature Selection Using ML-ReliefF and Neighborhood Mutual Information for Multilabel Neighborhood Decision Systems." *Information Sciences*, Vol 537, pp. 401-424. 2020.

[3] R. E. Schapire, Y. Singer, "Boostexter: a boosting-based system for text categorization." *Machine Learning*, Vol 39, No. 2-3, pp. 135-168, 2000.

[4] J. Read, L. Martino, P.M. Olmos, D. Luengo, "Scalable multi-output label prediction: From classifier chains to classifier trellises," *Pattern Recognition*, Vol 48, pp. 2096–

2109, 2012.

[5] Z. Barutcuoglu, R. E. Schapire, O. G. Troyanskaya, "Hierarchical multi-label prediction of gene function." *Bioinformatics*, Vol 22, No.7, pp. 830-836, 2006.

[6] A. C. E. Lima, L. N. De Castro, "A multi-label, semi-supervised classification approach applied to personality prediction in social media." *Neural Networks*, Vol 58, pp. 122-130, 2014.

[7] H. Xu, L. Xu, "Multi-label feature selection algorithm based on label pairwise ranking comparison transformation." *In 2017 International Joint Conference on Neural Networks (IJCNN)*, 2017.

[8] F. Li, D. Miao, W. Pedrycz, "Granular multi-label feature selection based on mutual information." *Pattern Recognition*, Vol 67, pp. 410-423, 2017.

[9] W. Chen, J. Yan, B. Zhang, Z. Chen, Q. Yang, "Document transformation for multi-label feature selection in text categorization." *In Seventh IEEE International Conference on Data Mining (ICDM 2007)*, 2007.

[10] A. Hashemi, M. B. Dowlatshahi, H. Nezamabadi-pour, "MFS-MCDM: Multi-label feature selection using multi-criteria decision making." *Knowledge-Based Systems*, Vol 206, 2020.

[11] M. Pushpa, S. Karpagavalli, "Multi-label classification: Problem transformation methods in Tamil phoneme classification." *Procedia Computer Science*, Vol 115, pp. 572-579, 2017.

[12] F. Luo, W. Guo, Y. Yu, G. Chen, "A multi-label classification algorithm based on kernel extreme learning machine." *Neurocomputing*, Vol 260, pp. 313-320, 2017.

[13] H. Xu, L. Xu, "Multi-label feature selection algorithm based on label pairwise ranking comparison transformation." *In 2017 International Joint Conference on Neural Networks (IJCNN)*, 2017.

[14] Y. Lin, Q. Hu, J. Liu, J. Chen, J. Duan, "Multi-label feature selection based on neighborhood mutual information." *Applied Soft Computing*, Vol 38, pp. 244-256, 2016.

[15] S. Kashef, H. Nezamabadi-pour, "A label-specific multi-label feature selection algorithm based on the Pareto dominance concept." *Pattern Recognition*, Vol 88, pp. 654-667, 2019.

[16] F. Li, D. Miao, W. Pedrycz, "Granular multi-label feature selection based on mutual information." *Pattern Recognition*, Vol 67, pp. 410-423, 2017.

[17] N. SpolaôR, E. A. Cherman, M. C. Monard, H. D. Lee, "A comparison of multi-label feature selection methods using the problem transformation approach." *Electronic Notes in Theoretical Computer Science*, Vol 292, pp. 135-151, 2013.

[18] J. Zhang, Z. Luo, C. Li, C. Zhou, S. Li, "Manifold regularized discriminative feature selection for multi-label learning." Pattern Recognition, Vol 95, pp. 136-150, 2019.

Table 2. Accuracy and Hamming Loss

		BR		CC		FW		Rakel		MCC		CDN		LC		Ensemble	
		Acc	**HL**	**Acc**	**HL**	**Acc**	**HL**	**Acc**	**HL**	**Acc**	**HL**	**Acc**	**HL**	**Acc**	**HL**	**Acc**	**HL**
BR Base	**CFS**	.620	.190	.673	.174	.654	.184	.654	.195	.673	.175	.596	.223	.641	.199	.706	.159
	IG	.549	.203	.634	.196	.628	.183	.541	.250	.634	.196	.547	.263	.606	.213	.704	.161
	ReliefF	.534	.210	.626	.201	.622	.186	.543	.249	.626	.201	.554	.257	.607	.216	.704	.163
	CV	.493	.232	.608	.223	.603	.202	.523	.275	.608	.223	.510	.291	.591	.23	.688	.178
CC Base	**CFS**	.282	.295	.593	.248	.590	.250	.591	.249	.554	.261	.378	.384	.587	.246	.599	.260
	IG	.549	.203	.634	.196	.628	.183	.541	.250	.634	.196	.547	.263	.606	.213	.704	.162
	ReliefF	.532	.208	.632	.195	.620	.185	.523	.258	.632	.195	.541	.268	.602	.214	.705	.160
	CV	.500	.229	.613	.219	.600	.201	.527	.269	.613	.219	.524	.284	.592	.229	.692	.175
LP Base	**CFS**	.634	.192	.674	.180	.659	.192	.667	.187	.674	.180	.570	.242	.646	.198	.688	.178
	IG	.628	.195	.672	.183	.663	.191	.667	.187	.665	.184	.566	.244	.644	.198	.688	.179
	ReliefF	.528	.211	.626	.201	.622	.184	.607	.197	.626	.201	.543	.267	.587	.219	.704	.162
	CV	.524	.211	.623	.201	.624	.186	.650	.209	.652	.197	.416	.331	.584	.221	.699	.165

Glass Shape and the Perception of Orange Juice

Chun-mao Wu, Hong-lei Zhou
College of Fashion and Design, Donghua University
Shanghai, China
[e-mail : cmwu@dhu.edu.cn]
*Corresponding author: Hong-lei Zhou

Abstract

Although previous research demonstrated the effects of glass shape on consumer perception and emotion, people have different sensory on orange juice which was influenced by containers. The main aim of this study was to investigate the effect of glass shape on the perceptions and on participants responses. This study was conducted in a lab where the orange juice was provided to participants (n = 30). The same quantity of the orange juice (100 ml) was offered to three types of glasses: a straight glass, a curved glass and a round glass. The data was collected through the questionnaire on intensity, sweetness, freshness, and pleasantness, which were rating by 7-point Likert scales. And the willingness to pay was also presented to participants. Compared to multisensory perceptions, the results prove that sweetness and pleasantness of the orange juice are influenced by glass shape. The findings show that the container is important to the juice consumption.

***Keywords*:** haptic; taste; visual; glass shape; multisensory

1. Introduction

Juice is the most popular beverage in the world, which affects health directly. It is important to understand perceptions of juice, and to test factors that influence sensory. The food perception has been affected by shapes and package materials (Schifferstein, 2009). Perception of drinks are influenced by the container (Hirson, Heymann, & Ebeler, 2012). There are some studies on the beer, cola, coffee, and other related alcoholic drinks, researchers proved the relationship between containers and drinks (Cavazzana, Larsson, Hoffmann, Hummel, & Haehner, 2017; Carvalho & Spence, 2018; Mirabito, Oliphant, Doorn, Watson, & Spence, 2017). The shapes of glasses can be adopted to intercept the public health (Troy, Maynard, Hickman, Attwood, & Munafò, 2015). In fact, the comparison was carried by researchers on two shapes of table glasses, one is short and wide glass, another is tall and slender glass, their findings confirmed that glass shape influence subjective responses (Cliceri, Petit, Garrel, Monteleone, & Giboreau, 2018). While, orange juice which preferred by consumers that seems not studied fully. The effects of the juice containers are important for consumers, designers and brand managers, which is influenced the perceptions and consuming experiences directly. To understand consumers' perceptions, this study was conducted to test the effects of different glass shapes.

1.1 Correspondence of shape and taste

Most of product experience associates with multisensory (Fenko & Rompay, 2018). The shapes of containers have proved that affect the consumption (Greenfield, Smith, & Wills, 1984;

This research was supported by a research grant from Shanghai Pujiang Program (No. 19PJC003)

Westerman, Gardner, Sutherland, White, Jordan, Watts, & Wells, 2012). Meanwhile, the food labels on the package influence consumer responses directly (Fenko et al., 2016). Adopting abstract shapes in the experiment, Spence and Ngo (2012) studied shape symbolic associations and proved that shapes on the package influence consumers' responses to the beverage. Plate colour and shape influence the taste perception in interacting experiences (Stewart, & Goss, 2013). In general, consumers' perception has some relationship with the shapes and materials. When the vessel is congruent to the drinks, people would like to drink more (Wan et al., 2015). In the early research, after the three conditions (visual, haptic, and visual+haptic), Balzartti, Maviglia, Biassoni, and Ciceri (2015) demonstrated that the glass is attached to higher quality perception than the plastic ones, furthermore, the perceived pleasantness is much more from the glass containers than the plastics. Given by this, the glass shape maybe an important factor in the consumption.

Compared to the previous findings in the laboratory contexts, glass shape influences consumers' behavior and responses in real-life context (Cliceri et al., 2018). Adopting three written scenarios (breakfast, movie and refreshing), the blackcurrant juices and apple juices are tested to evoke contexts, researchers find that pleasantness was enhanced by round glasses (Hein et al., 2012). Compared to the hedonic rating, the containers of the beverages have received far more investigation than the contexts. Multisensory is influenced by the container shapes. The perception is influenced by taste and shape features (Velasco et al., 2016). Furthermore, a number of studies have demonstrated that the shape of the glass influence people's feeling and experiences (Cavazzana et al., 2017; Mirabito et al., 2017). For example, people perceived more pleasantness and intensity in the typical glass (curved shape) than water glass (straight shape) or plastic bottle, that is congruent to the coca-cola (Cavazzana et al., 2017). Regarding the juice perception, it has been shown that the juice and glass shape influence the measurement. For example, the perceived pleasantness and intensity of juice were provided in glasses, which was interacted by visual, haptic, and taste. Through a shape-taste comparing study, Salgado-Montejo, Alvarado, Velasco, Salagado, Hasse, and Spence (2015) proved that participants judge the round and symmetrical shapes as pleasant and sweet, and the angular and asymmetrical is rated as unpleasant and sour.

The taste (bitter, salty, sour, sweet) and vision are affected by the shape features of angularity, asymmetry, and roundness in alcoholic drinks (Wan, Zhou, Woods, & Spence, 2015). Compared to the angularity and curved shapes, the sweetness is strongly correlated to the round shapes (Liang et al., 2013). The perceived quality is influenced by the containers' shape and the drinking colour. Meanwhile, the drinking experiences is related to the containers and perceptions. An empirical study on the judgement between coffee and cups, the taste attributes (i.e., sweetness, acidity) are significantly high on the split cup by the amateur and professional participants (Carvalho & Spence, 2018). The taste of coffee is influenced by shapes of containers, that the coffee in wider diameter mugs is associated to be sweeter, which is more aromatic in narrow diameter mugs (Doorn, Woods, Levitan, Wan, Velasco, Bernal-Torres & Spence, 2016). Round shapes are associated to the sweetness (Turoman, Velasco, Chen, Huang, & Spence, 2018; Velasco, Woods, Marks, Cheok, & Spence, 2016).

The shapes of the glass influence the drinking consumption. There are some studies that focused on the definition of freshness. For example, concerning the influence of the freshness, researchers demonstrated that perceived freshness is related to the healthy (Zhang et al., 2016). Both of the packaging attributes and the original products are contributed to the freshness perception (Péneau et al., 2009; Zhang et al., 2016). Especially, Attwood, Scott-Sameul, Stothart, and Munafo (2012) tested the relationship between the shape of glasses and rate of drinking, that the consumption behaviour is slower from a straight glass than a curved glass which is influenced by the effects of perceptual bias. Thus, compared to the straight glass, consumers tend to drink more beverage in the round shape of glass.

Compared to different shapes, the aroma of the coffee is stronger in the tulip cup than the split cup and the open cup (Carvalho & Spence, 2018). Shapes also influence the perception of intensity and fruitiness. For instance, after testing the straight sides and the curved sides, researchers found that the perceived fruitiness

and intensity are significantly influenced by the curved shape of glasses in beer drinking process (Mirabito et al., 2017). Perceived intensity is enhanced by round glasses (Hein, Hamid, Jaeger, & Delahunty, 2012). In the previous findings, both of curved and round shape which might be a factor that influence the perception of fruitiness and intensity. Furthermore, in the case study of the alcoholic drinks, researchers demonstrated that the willing to pay for the drinks would be increased as the content of the drinks are congruent to the shape of containers (Wan et al., 2015). Compared to the curved glass, the straight glasses are less monetary takings than the curved glass (Troy et al., 2015).

Although the shape and taste have been studied, the correspondence effects of shapes and perception still needs study. Further, the pleasantness combined with the consuming process empower it as an interesting topic to be studied.

1.2 Hypothesis

Usually, orange juice is served in the restaurant in the glass, the present study was designed for investigate whether different shapes of juice glasses influence the perceptual evaluation (e.g., aroma, sweetness, freshness, intensity, fruitiness, pleasantness, and willingness to pay) of the orange juice. The drinks are served in three juice glasses, as straight, curved or round shape of the glasses.

Consumers have different associations between the shape and beverage tastes. Researchers proved that people tend to associate round shapes to sweetness (Liang et al., 2013; Turoman et al., 2018; Velasco et al., 2016), as such, we hypothesized that participants would rate orange juice with the round side glass sweeter than that served in the straight glass or the curved glass. Similarly, people would perceive intensity that was served in curved or round glasses (Hein et al., 2012; Mirabito et al., 2017.). We hypothesis that participants would perceive more intensity which served in the round and curved glass than the juice served in the straight one. Furthermore, people perceived pleasant when they were served the juice in the round glass (Hein et al., 2012; Salgado-Montejo et al., 2015). We, therefore, expected that people would be more pleasant when they drink orange juice in the round glass than the juice served in the straight or curved glass. As mentioned already, people would like to pay more for the drinks in the curved containers (Troy et al., 2015). We expected that people would like to pay more for orange juice that served in the curved side glass.

2. Materials and Methods

2.1 Participants

There were 30 participants took part in this study (mean age = 22.5, SD = 2.48, ranging from 20 to 24 years). There participants were composed by 15 males and 15 females. The participants were recruited from the university, who would get a bottle of juice as credit. All the participants were consent before taking the experiment. The experiment was proved by the ethics committee at Donghua University. The testing was performed in the clean and quite room in a relax environment. All participants were naive to the experiment purpose and they were informed to sign the consent on the experiment.

2.2 Materials and Apparatus

Huiyuan Orange Juice was served to participants in the experiment. The juice contained 100% orange juice per 2 L box, no preservatives, colouring or artificial flavouring. The juice was adopted in the experiment that was stored in the refrigerator between the temperature 3 and 6 degree centigrade. The juice is described as fresh, transported in 0 to 5 degree centigrade, and no preservatives (Huiyuan Juice Group Cooperation, 2018). The participants were naive to the juice information so as to avoid bias.

The three glasses were chosen for this experiment. All of them were transparent and no colour (see **Fig. 1**). One glass was a juice glass with round sides (height: 18.1cm; diameter of opening: 8.5 cm; 473ml). The second one was characterized with straight sides (height: 16.7cm; diameter of opening: 6.2cm; 420ml). The third one was curved glass (height: 16.7cm; diameter of opening: 6.2cm; 420ml). The juices could be viewed through the transparent glasses.

Fig. 1. Three Different Glasses. (a)Round-shape Glass; (b) Straight-shape Glass; (c) Curved-shape Glass.

2.3 Design and Procedure

Experiment measures was adopted in this study to test the factors. A three kinds of cups (curved, straight, round) were used in the experiments. The variables are shown in Table 1. The taste perception was assessed by the variables of aroma, intensity, sweetness, freshness, fruitiness, pleasantness, and willingness to pay.

Three glasses (curved, straight, round) × one juice (orange juice) were used in the experiment. The 3 glasses were presented randomly. Each of them was filled with 100 ml orange juice. The participants were asked to drink the juice and rating the (1) aroma, (2) intensity, (3) sweetness, (4) freshness, (5) fruitiness, and (6) pleasantness. The variables were measured ranging from "1 = not at all" to "7 = extremely". The variable of willingness to pay was collected with an opening answer in the Chinese Yuan. At the end of the experiment, the participant would be asked the frequency that they drink juice (always, usually, often, sometimes, occasionally, or never).

The experiment was conducted in the Product Design and User Experience Lab, Donghua University. Before the experiment, the participants were asked if they were allergic to juice. If they said no, then the experiment was proceeded. The participants taken the experiment individually. The participants were informed to arrive the lab ten minutes earlier. A cup of water was provided to them.

The procedure followed one juices and

Table 1. Measurements

Dependent variable	Measures
Aroma	How aroma would you rate the perception of the juice?
Intensity	How intensity would you rate the perception of the juice?
Sweetness	How sweetness would you rate the perception of the juice?
Freshness	How freshness would you rate the perception of the juice?
Fruitiness	How fruitiness would you rate the perception of the juice?
Pleasantness	How pleasantness would you rate the perception of the juice?
Willingness to pay	How much would you like to pay for a glass of the juice?

three glasses (curved, straight, round) in this experiment. This study was conducted in a quiet room with 20 degree centigrade and the participants were sat in the places that were comfortable. Each glass was filled with 100 ml juices and kept in the refrigerator for 10 minutes. With the help of experimenter, the glass and questionnaire were provided to the participant, when the condition was completed, the next glass of orange juice and a questionnaire was provided to the participant. The order of the glass was randomized. In order to avoid the feeling mixture, participants were asked to sip some water and clean their mouth between drinking each juice. The time was recorded when the participants grasped the glass. In the visual-smell-taste condition, the participants were asked to taste the juice and rated the scales of aroma, intensity, sweetness, freshness, fruitiness, pleasantness, and willingness to pay.

After the experiment, the three glasses were rinsed by hot water and dried paper. They were kept in the refrigerator. The time totally no more than 15 minutes.

2.4 Analysis

The dependent variables were rated by seven-point Likert scale. The Analyses of Variance (ANOVAs) were conducted that to identify the dependent variables. The repeated-measures of glass shape (straight, curved, round) supported each ANOVA. products and group 2 (G2) with low familiarity with digital products. All the research process was completed in the participant's home through two weeks of in-depth observation. Meanwhile, the behavior and thinking of participants were recorded by video and questionnaire.

Table 2. Results of separate repeated-measures ANOVAs on dependent variables

Attribute	Outer-curved M	Outer SE	Straight M	Straight SE	Inner-curved M	Inner-curved SE	F	Sig.	Partial Eta Squared
Aroma	3.83	.25	3.77	.24	4.03	.21	.348	.707	0.008
Intensity	4.70	.23	4.37	.25	4.33	.22	.767	.468	0.017
Sweetness	3.20	.21	3.97	.24	4.33	.30	5.245	.007***	0.108
Freshness	4.27	.25	4.03	.22	4.63	.21	1.733	.183	0.038
Fruitiness	4.53	.23	4.30	.20	4.53	.22	.384	.682	0.009
Pleasantness	4.00	.19	4.43	.23	4.70	.25	2.512	.087*	0.055
Willingness to Pay	8.93	1.16	8.65	.91	9.20	.91	.076	.927	0.002

***P<0.01; **P<0.05; *P<0.1

3. Results and Discussion

The data was collected from 30 participants who follow the instruction of the experiment. The demographic character was collected. There were no missing values. The MCAR test was conducted so as to find if there was some missing data.

The shape of the glasses has significant effect on the attribute of 'sweetness', that was supported by the findings that $F = 2.512$, $p = 0.087, \eta_p^2 = 0.055$. The juice in the outer-curved glass (M=4.70) perceived [$F = 5.245$, $p = 0.007, \eta_p^2 = 0.108$]. The shape of glasses influenced the perception on sweetness of orange juice. The inner-curved glass (M = 4.33) was perceived sweeter than the straight (M = 3.97) and round ones (M = 3.20). This finding proved that inner-curved shape of glass was imaged to the perception of sweetness of juice. It is in line with the previous findings that the association affected the perception on containers and beverage.

The glass shape influences the association of 'pleasantness' significantly more pleasantness than straight sides (M = 4.43) and round sides (M = 4.00). The data analysis

was presented in Table 2. There are some association between the glass and drinks. Future study could research the novel shapes of juice containers. It is difficult to conclude that if there would be some association between sweetness and pleasantness on the juice and new containers.

4. Conclusion

This study explored if the perception on orange juice was influenced by the shape of glasses. The same volume of orange juice being provided in the different shapes of glasses, the association of sweetness and pleasantness were influenced by shapes. Therefore, based on the findings, the designers, manufactures and juice sellers should consider the glass shape and juices which are provided to consumers directly. In a realistic consumption, the taste maybe enhanced the haptic and visual. With the method of visual-haptic-taste, this study presented suggestions on the association between juice and glass shape, as it should be tested with beverages. Furthermore, these findings support the drinking experiences, test the variables in beverage consumption, and form a scale on the consumption evaluation.

References

[1] Attwood, A.S., Scott-Sameul, N.E., Stothart, G., & Munafo, M.R. (2012). Glass shape influence consumption rate for alcoholic beverages. *PloS One, 7(8),* e43007.

[2] Balzartti, S., Maviglia, B., Biassoni, F., & Ciceri, M.R. (2015). Glass vs. plastic: Affective judgments of food packages after visual and haptic exploration. *Procedia Manufacturing, 3,* 2251-2258.

[3] Cavazzana, A., Larsson, M., Hoffmann, E., Hummel, T., & Haehner, A. (2017). The vessel's shape influences the smell and taste of cola. *Food Quality and Preference, 59,* 8-13.

[4] Carvalho, F.M., & Spence, C. (2018). The shape of the cup influences aroma, taste, and hedonic judgements of specialty coffee. *Food Quality and Preference, 68,* 315-321.

[5] Cliceri, D., Petit, E., Garrel, C., Monteleone, E., & Giboreau, A. (2018). Effect of glass shape on subjective and behavioral consumer response in a real-life context of drinking consumption. *Food Quality and Preference, 64,* 187-191.

[6] Doorn, G.V., Woods, A., Levitan, C.A., Wan, X., Velasco, C., Bernal-Torres, C., & Spence, C. (2016). Does the shape of a cup influence coffee taste expectations? A cross-cultural, online study. *Food Quality and Preference. 56,* 201-211.

[7] Fenko, A., Kersten, L., & Bialkova, S. (2016). Overcoming consumer scepticism toward food labels: The role of multisensory experience. *Food Quality and Preference, 48,* 81-92.

[8] Fenko, A., & Rompay, T.V. (2018). Consumer-driven product design. *Methods in Consumer Research. Volume 2,* 427-462.

[9] Greenfield, H., Smith, A.M., & Wills, R.B. (1984). Influence of multi-holed shakers on salting on food. *Human Nutrition: Applied Nutrition, 38(3),* 199-201.

[10] Hein, K.A., Hamid, N., Jaeger, S.R., & Delahunty, C.M. (2012). Effects of evokated consumption contexts on hedonic ratings: A case study with two fruit beverages. *Food Quality and Preference, 26,* 35-44.

[11] Hirson, G.D., Heymann, H., & Ebeler, S.E. (2012). Equilibration time and glass shape effects on chemical and sensory properties of wine. *American Journal of Enology and Viticulture. 63(4),* 515-521.

[12] Liang, P., Roy, S., Chen, M.L., & Zhang, G.H. (2013). Visual influence of shapes and semantic familiarity on human sweet sensitivity. *Behavioural Brain Research, 253,* 42-47.

[13] Mirabito, A., Oliphant, M., Doorn, G.V., Watson, S., & Spence, C. (2017). Glass shape influences the flavour of beer. *Food Quality and Preference, 62,* 257-261.

[14] Péneau, S., Linke, A., Escher, F., & Nuessli, J. (2009). Freshness of fruits and vegetables: Consumer language and perception. *British Food Journal, 111(3),* 243-256.

[15] Salgado-Montejo, A., Alvarado, J.A., Velasco, C., Salagado, C.J., Hasse, K., & Spence, C. (2015). The sweetest thing: the influence of angularity, symmetry, and the number of elements on shape-valence and shape-taste matches. *Frontiers in Psychology. 6,* 1382.

[16] Schifferstein, H.N.J. (2009). The drinking

experience: Cup or content? *Food Quality and Preference, 20,* 268-276.

[17] Shanghai FS Food Company. Retrieved from http://www.fsjuice.com/.

[18] Stewart, P.C., & Goss, E. (2013). Plate shape and colour interact to influence taste and quality judgments. *Flavour, 2(27),*

[19] Spence, C., & Ngo, M.K. (2012). Assessing the shape symbolism of the taste, flavour, and texture of foods and beverages. *Flavour, 1,* 12.

[20] Troy, D.M., Maynard, O.M., Hickman, M., Attwood, A.S., & Munafò, M.R. (2015). The effect of glass shape on alcohol consumption in a naturalistic setting: a feasibility study. *Pilot and Feasibility Studies, 1,* 27.

[21] Turoman, N., Velasco, C., Chen, Y.C., Huang, P.C., & Spence, C. (2018). Symmetry and its role in the crossmodal correspondence between shape and taste. *Attention, Perception, & Psychophysics. 80(3),* 738-751.

[22] Velasco, C., Woods, A.T., Petit, O., Cheok, A.D., & Spence, C. (2016). Crossmodal correspondences between taste and shape, and their implications for product packaging: A review. *Food Quality and Preference, 52,* 17-26.

[23] Velasco, C., Woods, A.T., Marks, L.E., Cheok, A.D., & Spence, C. (2016). The semantic basis of taste-shape associations.*PeerJ, 4,* e1644.

[24] Wan, X., Zhou, X., Woods, A.T., & Spence, C. (2015). Influence of the glassware on the perception of alcoholic drinks. *Food Quality and Preference, 44,* 101-110.

[25] Westerman, S.J., Gardner, P.H., Sutherland, E.J., White, T., Jordan, K., Watts, D., & Wells, S. (2012). Product design: Preference for rounded versus angular design elements. *Psychology & Marketing, 29(8),* 595-605.

Research on the Influence of Cross-border E-commerce on Consumers' Purchasing Intention under the Internet+

Ziyang LIu, XiaoYin LIU, JiaYu LIU
Zhengzhou University, China.
[e-mail: morninglzy@hotmail.com]
*Corresponding author: Ziyang LIU

Abstract

The research purpose of this article is that the new consumption mode of purchasing foreign products through cross-border e-commerce platforms has some influence on consumers' purchasing intention. Product factors, social factors, trust factors, and efficiency factors are independent variables. Consumer purchase intentions use SPSS 21 statistical software and AMOS structural equation to verify the reliability, rationality, correlation analysis, and regression analysis of the assigned index data, and make assumptions Verification, the research results are: as the dependent variable, thus establishing the research model and hypothesis of this article.

***Keywords*:** Cross-border e-commerce, Consumer purchase intention,Perceived value.

1. Introduction

The establishment of Alibaba in 1999 marked the beginning of the development of cross-border e-commerce in China. Initially, Alibaba's Chinese suppliers were only yellow pages on the Internet, displaying product information of Chinese companies to global customers, positioning them in B2B bulk trade. in order to adapt to the increasing consumer demand, the study of cross-border e-commerce platforms' influence factors on consumers' shopping intentions will be conducive to the construction and improvement of cross-border e-commerce platforms, promote the development of cross-border e-commerce enterprises, and promote the development of cross-border e-commerce companies [1]. The development of foreign trade is of great help.

2. Research model establishment and research hypothesis

2.1 Operational definition

Product elements refer to the measurement items of the eTransQual scale developed by Bauer (2006) based on the consumer's online shopping process; social elements and trust elements refer to Schmidt (1999) [2.3]. This experience is manifested in the trust in the B2C platform of cross-border e-commerce and consumer social psychology: Research shows that this kind of customer trust will influence customers' online consumption decisions through psychological effects; the measurement items of efficiency attributes refer to Lee (2010) Zhang (2014) [4.5]; Perceived value reference Vasilis (2005) divides perceived value into three factors, professional services, marketing efficiency, and effective communication. Consumer purchase

expectations refer to Petrick & Dubinsky (2003) and Park et al. (2007) [6.7]) Scale measurement items.

2.2 Research model

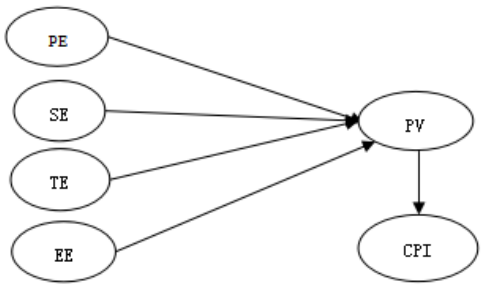


Fig. 1. Research model

H1: Product elements have a significant positive effect on perceived value

H2: Social factors have a significant positive impact on perceived value

H3: Trust factors have a significant positive effect on perceived value

H4: Efficiency factors have a significant positive impact on perceived value

H5: Perceived value has a significant positive effect on consumer purchase intention

3. Research methods and research results

3.1 Research method

According to the research objectives of this article and the specific characteristics of the research objects, a sampling survey will be used. The survey time is from September 5, 2019 to October 10, 2019. First, prepare to collect 100 questionnaires, hoping to find out the problems in the questionnaire through the feedback of 100 consumers, so as to improve the revision of the questionnaire, and then implement the formal questionnaire survey and distribution. A total of 424 questionnaires were sent out. After the questionnaires were collected, the data was checked by Excel, which included a total of 10 invalid questionnaires. In the end, 414 valid questionnaires were actually collected, and 98% of them were actually collected. Finally, 414 valid questionnaires were used for analysis, and SPSS 22.0 was used for reliability analysis, factor analysis and structural equation model analysis.

3.2Research result

In terms of the verification of basic adaptation indicators of the model, according to the verification method of scholars such as Bogozzi (1988), as shown in Figure 2. Chi-square value = 432.103, GFI = 0.883, AGFI = 0.838, RMSEA <1 and other results can be obtained. According to the indicator criterion, we can completely determine that this result can pass the model fitness criterion.

H1: The influence of product elements on perceived value is not established

H2: Social factors have a significant positive impact on perceived value

H3: Trust factors have a significant positive effect on perceived value

H4: Efficiency factors have a significant positive impact on perceived value

H5: Perceived value has a significant positive effect on consumer purchase intention

4. Analysis conclusion

Through empirical research, the following indicators are obtained: social elements, trust elements, and efficiency elements in cross-border e-commerce have a positive impact on consumers' perceived value, and perceived value has a positive impact on consumer purchases. First of all, it shows that the future development of cross-border e-commerce should combine online and models to enhance consumers' experience of product perception factors and accelerate consumers' decision to purchase

intentions.Second, from data analysis, social factors, trust factors, and efficiency factors have the greatest impact on consumers' perceived value. It is suggested that cross-border e-commerce platforms should pay attention to consumers' evaluations of products, actively obtain consumer feedback in terms of after-sales service, and realize the concept of customer service value management.At present, for domestic cross-border e-commerce platforms, the logistics information of consumers purchasing products can only be inquired after the products enter customs, thus attaching overseas logistics to the products On the home page, direct mail products can be connected with overseas logistics providers, and consumers can inquire about overseas logistics information.

References

[1] Sun Lei, Wang Fang. China's cross-border e-commerce development status and countermeasures[J]. China Circulation Economy, 2015 (3): 38-39

[2] Xinyue. Who controls the measurement of cross-border e-commerce [J]. Business Culture, 2015 (11): 49-57.

[3] Liao Zhong, Zhang Maoteng. An empirical study of Internet-based e-shopping and consumer attitudes[M]. 2001.

[4] Slyke C V, Belanger F, Sridhar V. Comparison of American and Indian consumers' views on e-commerce [M] //Comparison of American and Indian consumers' views on e-commerce. 2005: 24-40.

[5] Chen Y H, Hsu I C, Lin C C. Website attributes to improve consumers' purchasing intention: joint analysis [J]. Journal of Northeastern University (Natural Science Edition), 2009, 28 (2): 155-159 Journal of Business Research, 2010, 63 (9): 1007-1014.

[6] Law R, BaiB. How are the preferences of online buyers and browsers different in the design and content of travel websites? International Journal of Contemporary Hotel Management, 2008, 20(4): 388-400.

[7] Zeithaml V A. Consumers' perceptions of price, quality and value: a mean-end model and evidence synthesis[J]. Economic Survey, 2013, 28(1): 109-114. Journal of Marketing, 1988, 52(3): 2-22.

Unsupervised Image-to-Image Translation Using Generative Adversarial Networks Based on Dual Attention Network

Yulin Yang and Rize Jin
Department of Computer Science, Tiangong University
Tianjin, China
[yangyulin1991@gmail.com, jinrize@tiangong.edu.cn]
*Corresponding author: Rize Jin

Abstract

The state-of-the-art GANs-based image translation models can perform image translations for multiple domains with a single generative model. However, these methods are excellent in convert low-level information through the use of reconstruction error, but fail to make a reasonable translation when translating images on a specific task, such as, editing gender of a face image. To address these limitations, we propose the generative adversarial networks based on dual attention network and perceptual loss for unsupervised image-to-image translation, dual attention network not only allows attention-driven, long-range dependency modeling for image translation in multiple domains, but by exploiting the interdependencies between channel maps, the proposed model could emphasize interdependent feature maps and improve the quality of image translation. Furthermore, recent works have shown that the perceptual loss that utilizing the image features extracting from a layer of pre-trained model to guide image translation affects the performance of image translation. A comprehensive experiment is designed for studying the effects of the perceptual loss on the image translation. Experimental results demonstrate that the proposed model can show better result than the previous models on the quality of generated image. The proposed model performs better than comparative work, reducing Kernel Inception Distance (KID) and Fréchet Inception Distance (FID) from (11.37, 100.15) to (10.53, 92.60) respectively on the challenging dataset.

***Keywords*:** Image Translation, Dual Attention Network, Perceptual Loss

1. Introduction

The emergence of the term called Generative Adversarial Networks (GANs)[1] on 2014 introduced by the Goodfellow gives rise to crazy of the researcher among multiple communities, especially in the field of computer vision. GANs are now ubiquitous and applied to image editing, data generating, predict attention and 3D subject generating, recently, even generate awe-inspiring artwork: Image-to-Image and Style translation.

There are lots of studies for doing image translation tasks. Among them are two main methods: image translation based on supervised learning and unsupervised learning. Phillip Isola, et al. proposed a promising model of their Pix2Pix method in [2] based conditional generative adversarial network to translate image. It not only learns the mapping from input image to result, but also learn a loss function to train this mapping. This way is effective at synthesizing photos from label maps, reconstructing objects from edge maps, and colorizing images such as translate semantic images to photographs of

cityscapes and builds, photographs to Google Maps, and white photographs to color, sketches to color photographs and so on. However, the model of Pix2Pix requires a set of image pairs for training. For many tasks, it is difficult to prepare paired training data, and sometimes the paired data is not available. Subsequently, Cycle-GAN model was proposed in [3] by Jun-Yan Zhu on 2017. In order to learn mapping function between unpaired datasets, it introduces a novel loss function called cycle consistency loss. It gets compelling performance on several tasks such as style transfer, object transfiguration, season transfer etc. But, Cycle-GAN still failed to some cases and has some limitations particularly in geometric changes. Although These Cycle-GAN have remarkable success image-to-image translation on unpaired datasets, they are short of scalability and robustness in more than two domains that mean that we have to build a model separately for every domain, which is cumbersome. A scalable model called Star-GAN is proposed by Yunjey Choi et al. in [4] which can perform image-to-image translation for multiple domains using only a single model by taking a domain label as an additional input to generator of the model. However, it failed to capture the multi-model nature of the data distribution. YunJey Choi et al.in [5] improved the StarGAN model by adding a mapping network to original StarGAN. The improved StarGAN called StarGANv2 can generate diverse images across multiple domains in high resolution images. However, by carefully examining the generated samples from these generative models, we can observe that even the state-of-the-art models (Star-GAN) has some problems on image translation, for instance, the background of the generated image is easily affected and the result of generated image is unreasonable.

Motivated by human attention mechanism theories, attention mechanism has been successfully introduced in computer vision tasks, for example, image-to-image [7], image generation [8], image segmentation [9]. Rather than compressing an entire image or a sequence into a static representation, attention allows the model to focus on the most relevant part of images, which is conducive to exhibits a better balance between the ability to model long-range dependencies to a certain degree. [8] propose the Self-Attention Generative Adversarial Network (SAGAN) which introduce a self-attention mechanism into convolutional GANs. Armed with the self-attention, the generator can draw images in which fine details at every location are carefully coordinated with fine details in distant portions of the image. [7] propose the Attention-Guided Generative Adversarial Networks (AGGAN) for unsupervised image-to-image translation, which produce attention masks via a built-in attention mechanism, and then fuse the input image with the attention mask to achieve image translation with high-quality. [9] propose a Dual Attention Network (DANet) for Scene Segmentation, which can adaptively integrate local features with their global dependencies.

2. Related Works

2.1 Generative Adversarial Network

The GAN is a type of neural network architecture for generative model which consists of two sub-networks Generator G that captures the data distribution, and Discriminator D that estimates the probability that a sample came from the training data rather than Generator G. The model is trained in adversarial way for G is to maximize the probability of D making a mistake, while two sub networks contest with each other in the sense of game theory.

$$\frac{min}{G}\frac{max}{D}E_{x\sim p_{data}}\left[\log\left(D(x)\right)\right]+E_{z\sim p_z}[\log(1-D(G(z)))] \quad (1)$$

GAN is also famous for its unstable training and mode collapse. Subsequently, WGAN [10] and WGAN-GP [11] are proposed to solve this problem. Both of them have proved to be very effective.

2.2 Unsupervised Image-to-Image Translation

Unsupervised image-to-image translation is an important application of GANs, which means to map an image from one domain to another or multiple domain without paired datasets. Cycle-GAN proposed the cycle consistency loss to enforce the learning process that if an image is translated to the target domain by learning a

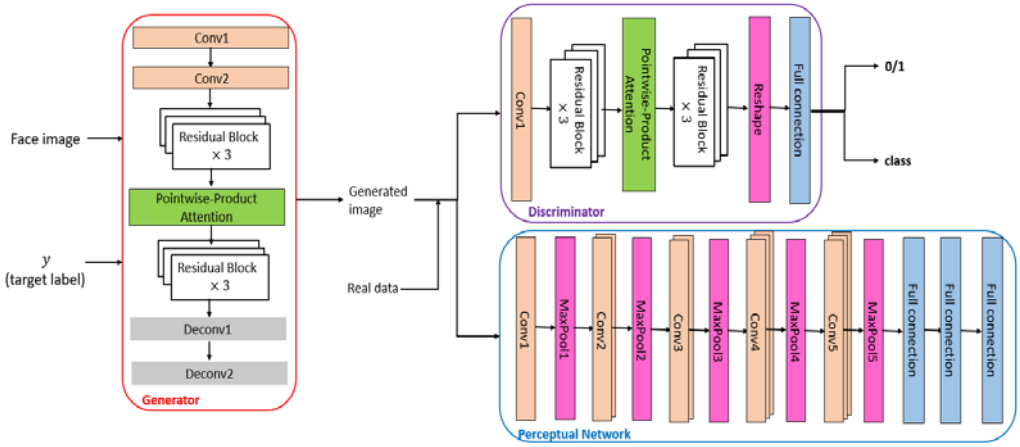


Fig. 1. The Broad Architecture overview of the generator, discriminator and perceptron framework presented in this paper.

mapping and after an inverse mapping, the output should be same with the original image, $X = G(G(X))$. Star-GAN provides a remarkable benchmark architecture for multi-domain image-to-image translation, which introduce classification loss for the image translation between different domains. Although some image translation models have been proposed, most of them are built on Star-GAN. Therefore, this paper mainly compares with Star-GAN.

3. Framework

3.1 Loss Function of Proposed Model

As shown in **Fig. 1**, our attention-GAN is comprised of three basic subnetworks, i.e., a generative networks G with dual attention mechanism that is shown in [9] and a discriminator D with an auxiliary classifier and a per-trained VGG-19 network. Given face image X and target domain c, the generative networks G is used to translate an input image X into an generated image X_{fake} according with target domain c. In order to enable the generator to generate a new face with target domain c while preserving other details. On the one hand, we introduce a classification loss in our model, which ensured that the generated image X_{fake} must have the desired attributes, on the other hand, we also introduce a reconstruction loss in our model, which constrain the generated image X_{fake} must be consistent with the input image x as possible.

Adversarial Loss: Instead of use loss function of vanilla GAN, the WGAN-GP [11] is used to train adversarial loss to avoid mode collapse. The generator G generates an image X_{fake} conditioned on both the input image X and the target domain c, which is shown in equation 2.

$$L_{adv} = E_x[D_{src}(x)] - E_{x,c}\left[D_{src}\left(G(x,c)\right)\right]. \quad (2)$$

In this paper, we define $D_{src}(\mathrm{x})$ as a probability distribution of the source data given by D.

Classification Loss: Aim to make the category of generated image correspond to the target label c, an auxiliary classifier is added on top of the discriminator D to control the category of generated image. The process is defined as Equation.3.

$$L_{cls} = -\mathrm{E}_x\log([\mathrm{D}_{cls}\ (\mathrm{c}\tilde{}|\mathrm{x})] - E_{x,c}\left[\log\left(D_{cls}\left(c|G(x,c)\right)\right)\right] \quad (3)$$

where the term $D_{cls}(c|\mathrm{x})$ represents a probability that the input image X come from c domain, which is computed by D.

Reconstruction Loss: In order to guarantee that the generated image can change only the domain-related part of the inputs while preserve the other content of its input image. we introduce a cycle consistency loss [3] into the generator, which is formulated as Equation. 4:

$$L_{rec} = \mathrm{E}_{\mathrm{x}\sim \mathrm{p}_{data}, x_{fake}\sim p_g}[|X - G(X_{fake}, c^{\sim})|] \quad (4)$$

Perceptual Loss: The perceptual loss function has been introduced into [12] for style transfer and super-resolution. Instead of relying only on L1 or L2 loss on real image and generated image, The VGG16 [13] model trained on ImageNet is used to extract semantic features, and the L1 distance between these features of real and fake images is used to guide the generator G. The perceptual loss function is defined as:

$$L_{pec} = ||V_i(X) - V_i(\mathrm{X}_{fake})||_1 \quad (5)$$

Overall Objective: Finally, the objective for our proposed model is formulated as Equation 4 and Equation 5, which are used to optimize G and D respectively.

$$L_D = \mu_1 * \mathrm{L}_{adv} + \mu_2 * L_{cls}^D$$
$$L_G = \mu_3 L_{cls} + \mu_4 * L_{rec} + \mu_5 * L_{pec}$$

from μ_1 and μ_5 represent weight of adversarial loss, classification loss of D, classification loss of G, and reconstruction loss and semantic perceptual loss respectively. In this work, all weights are set to 1.

3.2 Architecture

As is shown in **Fig. 1**, the generator of proposed model mainly contains the following four parts: which are convolution layers, dilated convolution layers, deconvolution layers and dual attention model. Instead of using standard convolution layers to encode the input images into a low dimensional vector, dilated convolution layers is used for enlarging receptive fields. And then dual attention model is used to enhance the detailed feature extraction and aggregates long-range contextual information of the generator. Beyond that, two convolution layers and two deconvolution layers are added to the front end and back end of dual attention operations, respectively. Each layer of convolution layers, dilated convolution layers and deconvolution layers mainly includes following four operations: convolution, instance normalization (IN), Rectified Linear Unit (ReLU). The kernel sizes of convolution layers are 7 × 7, 4 × 4. The kernel sizes of all of dilated convolution layers are 3 × 3. The kernel sizes of are 4 × 4, 7 × 7. The numbers of channel are 64, 128, 256, 128, 64, 64, 3, respectively.

For discriminator, as shown in Fig.1, the CNN-based discriminator only consists of eight convolution layers and a dual attention model. More specifically, the network of discriminator are seven convolution layers with 4 × 4 kernel sizes. And then a dual attention model is used to improve the ability of discriminator. Finally, two parallel convolution layers, one for judging the input image come from real data or generated image, the other for judging the categories of input images.

4. Experiments

Dataset: To evaluate the proposed method, some experiments are conducted on CelebA-HQ [17]. Compared with the CelebFaces Attributes (CelebA) database, CelebA-HQ is a rather challenging dataset, which has a total of 30K images, 40 binary attributes annotations per image and each with a resolution of 1024 * 1024. In this work, we randomly selected 2,000 and 1,000 images from CelebA and CelebA-HQ for testing and quantitative evaluating, the rest used for training, respectively.

Evaluation Metrics: We choose two well-known metrics for quantitative evaluation: Kernel Inception Distance (KID) [14] and Fré chet Inception Distance (FID) [15]. Though alternatives exist the Inception score (IS) [16] and Amazon Mechanical Turk (AMT), but they are either inappropriate or costly. The Inception score computes the KL divergence between the conditional class distribution and the marginal class distribution. However, it has serious limitations—it is intended primarily to ensure that the model generates samples that can be confidently recognized as belonging to a specific class, and that the model generates samples from many classes, not necessarily to assess realism of details or intra-class diversity. FID and KID are more principled and comprehensive metric, and has been shown to be more consistent with human evaluation in assessing the realism and variation of the generated samples. The former calculates the Wasserstein-2 distance between the generated images and the real images in the feature space of an Inception-v3 network. the latter Measures the dissimilarity between two probability distributions the real and the generated samples using samples drawn independently from each distribution. Lower values of them mean closer distances between synthetic and real data distributions. In all our

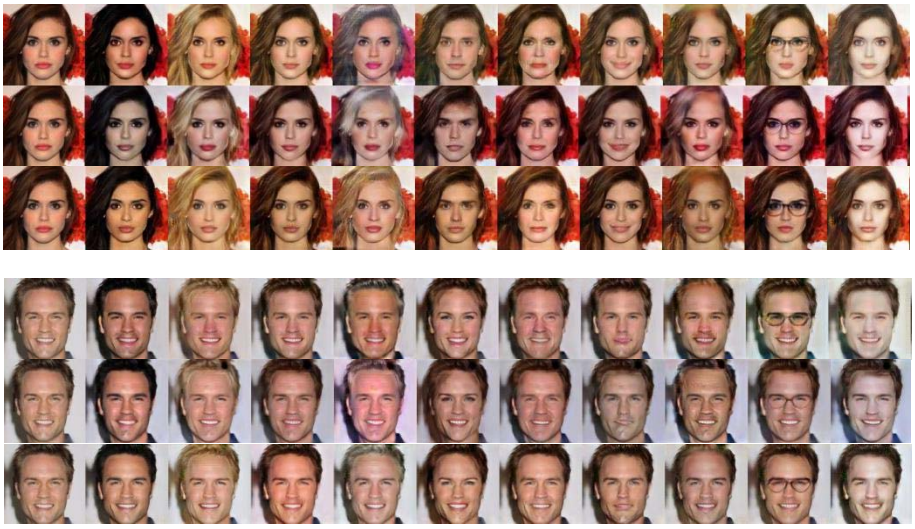


Fig. 2. Comparison with different baselines on celebA datasets and celebA-HQ datasets, from left to right are source domain image, 'Black Hair', 'Blond Hair', 'Brown Hair', 'Gray Hair', 'Male', 'Young', 'Smiling', 'Bald', 'Eyeglasses', 'Pale Skin', respectively. From top to bottom, each three rows were grouped in pairs. The first, second, and third rows are star-GAN, DA-GAN, and DA-GAN with perceptual losses respectively in each grouped.

	Star-GAN		DA-GAN		DA-GAN + Ploss	
	Blond Hair	Gender	Blond Hair	Gender	Blond Hair	Gender
FID	100.15	28.32	96.39	30.99	**92.60**	**25.99**
KID	11.37	12.40	10.62	15.83	**10.53**	**10.15**

Table 1. Quantitative comparison with different models in different domains. On the left is the score of editing hair color, on the right is the score of editing gender.

experiments, 2k samples are randomly generated for each model, the same number samples for the specified attributes are generated for each model.

4.1 Multiple Facial Attribute Editing

In this section, we compare the performance of star-GAN, DA-GAN, DA-GAN with perceptual loss in terms of multiple facial attribute editing. Specifically speaking is ten attributes with strong visual impact in CelebA-HQ datasets are chosen in our experiments, which contain 'Black Hair', 'Blond Hair', 'Brown Hair', 'Gray Hair', 'Male', 'Young', 'Smiling', 'Bald', 'Eyeglasses', 'Pale Skin'. As shown in **Fig. 2**, we can see the following phenomena: First, the quality of image generation from star-GAN is a little fuzzy in comparison to the proposed models, one reasonable explanation for such cases could be that two attention mechanisms are introduced into the proposed models; One is the GAM model that captures global dependencies regardless of locations, which adaptively aggregate long-range contextual information to make the framework context-aware, the other is CAM that emphasizes interdependent feature maps and improve the feature representation of image translation. Then, although the quality of the image generation is very satisfactory, we're detecting some unexpected experimental results. For instance, when modifying the following attributes: 'Male', 'Young', 'Smiling', 'Bald', 'Eyeglasses', 'Pale Skin', the Hair color of input image have changed considerably, which was not to be expected. About this problem, we hypothesized that previous image translation models can perform image translations for multiple domains with a single generative model by imposing class constraints on generator. However, when the constraints increase, the generator will ignore some information to meet the above increasing constraints, which lead to the generator is hard to focus on desired attribute. namely the generator will generate a new face with desired attributes, but fail to preserve other details well.

In order to evaluate our model quantitatively, KID and FID is used, the results are shown in

Table 1. Compared with Star-GAN, DA-GAN is more effective than Star-GAN in editing Hair color. DA-GAN with perceptual loss Got the best performance in editing Hair color and gender, which further shows that our model is effective.

5. Conclusion and Future Work

This paper proposed model called generative adversarial networks based on dual attention network and perceptual loss for unsupervised image-to-image translation. Based on these results we could say that proposed method not only improve the quality of image translation, but also have a great performance for more reasonable image translation. The task of high-resolution image translation has always been a research hotspot, we will attempt to solve this problem in the next stage.

References

[1] Goodfellow, Ian, et al. "Generative Adversarial Nets." Advances in Neural Information Processing Systems 27, 2014, pp. 2672–2680.

[2] Isola P, Zhu J, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. In: 2017 IEEE conference on computer vision and pattern recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017; 2017. p. 5967–76.

[3] J. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-toimage translation using cycle-consistent adversarial networks," in IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017, pp. 2242–2251, 2017.

[4] Choi, Yunjey, et al. "StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation." 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 8789–8797.

[5] Choi, Yunjey, et al. "StarGAN v2: Diverse Image Synthesis for Multiple Domains." 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 8188–8197.

[6] Liu, Ziwei, et al. "Deep Learning Face Attributes in the Wild." 2015 IEEE International Conference on Computer Vision (ICCV), 2015, pp. 3730–3738.

[7] Zhang, Han, et al. "Self-Attention Generative Adversarial Networks." ICML 2019 : Thirty-Sixth International Conference on Machine Learning, 2019, pp. 7354–7363.

[8] Fu, Jun, et al. "Dual Attention Network for Scene Segmentation." 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 3146–3154.

[9] Hu, Jie, et al. "Squeeze-and-Excitation Networks." 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 7132–7141.

[10] Arjovsky, Martin, et al. Wasserstein GAN. 2017.

[11] Gulrajani, Ishaan, et al. "Improved Training of Wasserstein GANs." NIPS'17 Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017, pp. 5769–5779.

[12] Johnson, Justin, et al. "Perceptual Losses for Real-Time Style Transfer and Super-Resolution." European Conference on Computer Vision (ECCV), 2016, pp. 694–711.

[13] Simonyan, Karen, and Andrew Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition." ICLR 2015 : International Conference on Learning Representations 2015, 2015.

[14] Bińkowski, Mikołaj, et al. "Demystifying MMD GANs." ICLR 2018 : International Conference on Learning Representations 2018, 2018.

[15] Mathiasen, Alexander, and Frederik Hvilshøj. "Fast Fréchet Inception Distance." ArXiv Preprint ArXiv:2009.14075, 2020.

[16] Barratt, Shane T., and Rishi Sharma. "A Note on the Inception Score." ArXiv Preprint ArXiv:1801.01973, 2018.

[17] T. Karras, T. Aila, S. Laine, and J. Lehtinen. Progressive growing of GANs for improved quality, stability, and variation. CoRR, abs/1710.10196, 2017.

KSII The 12th International Conference on Internet (ICONI) 2020.

MVC based Approach to Consume Weather API: A Mobile Phone System

Maria Clara Gracia, Cassandra Lefever, Dana Long, Franklin May, Imran Ghani
Mathematical and Computer Science Department
Indiana University of Pennsylvania
E-mail: mjjz@iup.edu, gkgy@iup.edu, cnvw@iup.edu, wyxy@iup.edu, ighani@iupe.du
*Corresponding author: Franklin May

Abstract

Weather forecasting started being implemented in 1861 when it was published in the Times newspaper. This idea was introduced by Admiral Robert FitzRoy. This idea was created to provide readers with an accurate prediction of the weather to influence their daily lives. Nowadays, weather is most consumed through either weather applications or television broadcasts. Weather applications commonly have radars, news articles, and forecasts. Weather applications use information obtained through weather satellite imagery to publicly distribute it through a user-friendly interface. However, many of these applications fail to be user friendly because of cluttered interfaces and lack of notification control. This may include too many pictures or articles that take away from the essential information that the users would like to see when opening the application. Our software, WeatherTime, is an application that provides its users with current and past weather reports. It will be accessed through a user-friendly mobile app. The data gathered for our application will be accumulated using an API to provide the most accurate information possible. It will send casual and fun notifications to keep the users engaged and focus on informing them of how the weather will impact them. WeatherTime will not present inaccurate information or overwhelm its users with a plethora of unnecessary materials.

Keywords: Weather forecast; Application programming interface (API); Interface; Notifications

1. Introduction

WeatherTime is a weather application that will provide its users with current and past weather reports. This idea was introduced by Admiral Robert FitzRoy [1]. In WeatherTime it will be possible to see the weather for the next 10 days as well as a weather radar, and notifications for the weather. WeatherTime will have two categories of users. There will be a base user and then a premium user. The base user will be able to access the basic weather. Then the premium user will be able to access their payment records, radar data, and past weather records.

2. Model-View-Controller Approach to Develop WeatherTime System

In this project, we implemented a MVC design pattern [2]. **Fig. 1.** We present the package diagram of our MVC design pattern with 3 packages in it.

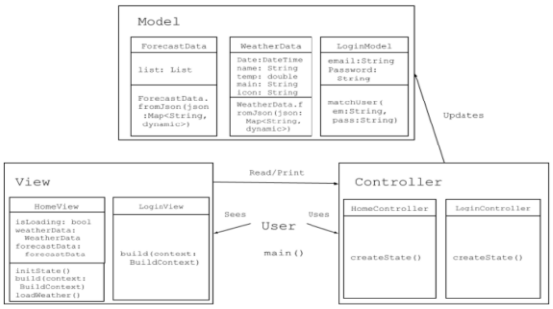


Fig. 1. MVC package diagram

2.1 Model Component

The model component is the area where the data-related logic is stored for the user to access. In the WeatherTime app, this is where the forecast, weather, and login data are stored. The weather data consists of information such as the current temperature, the location name, the date, and the weather icon. The forecast data consists of the predicted temperature for future dates. The login data consists of the email and the password. We called OpenWeatherAPI [3].

2.2 View Component

The view component is the area that the end user interacts with. For the WeatherTime app, this consists of the home view and the login view. The home view provides the GUI for the home page that the user sees upon opening the app. The login view is the page the user sees when logging into this application.

2.2 Controller Component

The control component is the area that passing data between the model and the view. In the WeatherTime app, this is the login controller and the home controller. It takes the users input into the GUI and transfers that data over to the model to process a login.

3. System User Interface

When the user opens the WeatherTime App, **Fig. 2** will be displayed to provide an overview of the weather to the user.

Fig. 2. GUI of Dashboard

4. Kanban-based Sprint Planning using Trello

We used Kanban board using Trello software as shown in **Fig. 3.** We created three (3) sprints for this class project in each sprint we divided our tasks between the members of our team and used Trello for documenting what tasks were under the process, what tasks were finished, and what tasks we had yet to start. We determined what tasks were going to be developed on each sprint based on the importance of it to the overall app and dependencies between tasks and determined if any changes needed to be made. After all of the tasks in a sprint were completed, we tested all of the components and determined if any changes needed to be done to the application or if any errors were going to be moved to the next sprint.

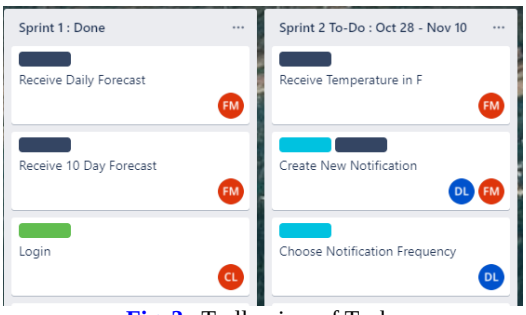


Fig. 3. Trello view of Tasks

5. Conclusions

In this paper, we present a mobile application for accessing accurate weather information using OpenWeatherAPI for accurate and up to date weather information. Our mobile application contains a basic GUI with visual aids supported by OpenWeatherAPI to give the user a brief overview of the forecast for the day. The development team of WeatherTime used Trello as seen in **Fig. 3**. to properly distribute tasks in order to complete production on the application of WeatherTime in a timely manner. Improvements to WeatherTime will continue to be developed based on input from the users in order to compete with other major Weather applications.

References

[1] https://en.wikipedia.org/wiki/Model%E2%80%93view%E2%80%93controller
[2] Brown, B. R. "The First Weather Forecast, August 1, 1861", https://bit.ly/3pkSm46, 2011.
[3] Weather API - OpenWeatherMapopenweathermap.org

MVC Based Approach to Implement Employees Time Tracking System

Rohith Gattu, Shane Peterson. Tyler Omune, Logan Peters, Imran Ghani
E-mail : xcfy@iup.edu, pnfy@iup.edu, hyhbc@iup.edu, yhgy@iup.edu, ighani@iup.edu
Department of Mathematical and Computer Sciences
Indiana University of Pennsylvania
1011 South Drive
Indiana, PA 15705
*Corresponding Author: Rohit Gattu

Abstract

Employee Management Systems has been an integral part of complex businesses for years due to its efficiency for day-to-day tasks, employee hours, and other miscellaneous information. On Da Clock is aimed to efficiently and mindfully track employee's hours as well as provide daily details within the specified company. Nowadays, clock-in systems have become a hub for company information as well as payroll. Traditional payroll systems offer a simple Graphical User Interface (GUI) while upholding sensitive information to the highest security standard. Communication programs such as Microsoft Teams provide information and intercommunication within a division, company, and/or branch. With the combination of both systems, companies' efficiency and organization can be accumulated and built upon to improve the virtual environment. The implementation of the Model-View-Controller (MVC) based approach dissembles our program into specified areas, which provides the program to become more robust and can be tailored to the customers liking more easily. This type of software's architecture is based heavily on the three-layer architecture by applying the component-based approach compared to other similar programs. However, the component approach provides a small scale of reusable features. Many clock-in systems do not have capabilities to facilitate communication as well as payroll. Still, when using the MVC-based approach, we can focus specifically on features that a company intends to use, creating a personalized experience for each component if necessary and keeping a component generic. This paper proposes to provide a program that facilitates an MVC based approach which offers a wide range of business to an all-in-one intercommunication, employee hours, and micaceous information system which can be tailored to the customers' requirements.

***Keywords*:** Employee Management System, Model-View-Controller (MVC), Time Management, Office Communication System

1. Introduction

With the increase of multi-user infrastructures in business, employee management systems have become a crucial communication standard. The demand for such a system has increased exponentially in many leading industries; however, with specific needs and requirements for a multitude of companies, there has been a new approach to handling such programs. The MVC-based approach focuses specifically on features that the company wants specifically while creating a personalized experience for the components of the program they select [1]. With this approach, we are able to develop a robust generic model as a basis for a business to customize the program further. This provides the development team efficient progress when companies plan the customization of the software.

2. Model-View-Controller Approach to Develop the System

In this project, we decided to use a model-view-controller approach to utilize the software components better. **Fig. 1** shows the package diagram with the further improvement of adding an MVC-based approach.

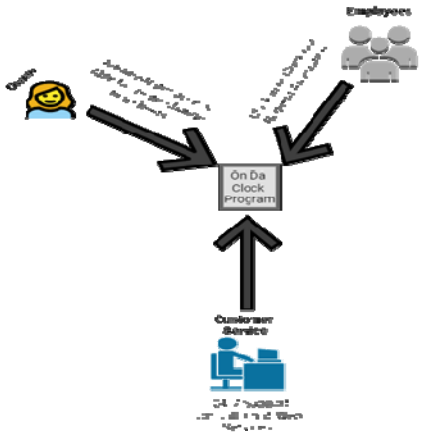


Fig. 1. MVC package diagram

2.1 Model Component

The model shows three main aspects of the programming providing the makeup of the software itself. With these three components, we are able to treat them as separate pieces of software with various custom components.

2.2 View Component

The GUI can be used in various ways to access the entire system. With the implementation of Role-Based Access Control (RBAC), we can secure the system when a user utilizes the view component of the program [2].

2.2 Controller Component

Controlling the system can be accessed via the company's administration or technical support, which can troubleshoot remotely due to the secure end-to-end encryption [3]. The controller component provides versatility as well as the efficiency of the internals of the program.

3. System User Interface

The system provides users with a simple yet intuitive GUI that shows vital information that the user will likely interact with often. **Fig. 2** shows the login screen of the program with options of resting the users password as well as creating anew user within the company.

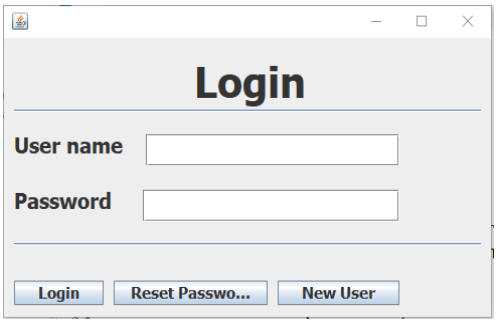


Fig. 2. GUI of Dashboard

4. Kanban-based Sprint Planning using Trello

For this project's organization, we used Kanban-based planning with the help of Trello, a web-based list-making application that provides useful information on tasks shown in **Fig. 3.**

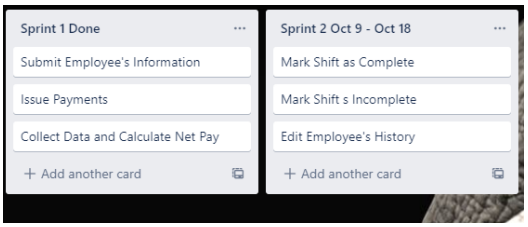


Fig. 3. Trello view of Tasks

5. Conclusions

This paper proposes to provide a program that facilitates an MVC based approach which offers a wide range of business to an all-in-one intercommunication, employee hours, and micaceous information system

References

[1] A. Hornsby, "XMPP message-based MVC architecture for event-driven real-time interactive applications,",pp. 617-618, doi: 10.1109/ICCE.2011.5722771.

[2] Ferraiolo, David & Kuhn, D.. (2009). Role-Based Access Controls.

[3] Bai, Wei & Pearson, Michael & Kelley, Patrick & Mazurek, Michelle. (2020). Improving Non-Experts' Understanding of End-to-End Encryption: An Exploratory Study.

MVC Based Approach to Implement Prediction and Betting System

Hunter M. Sargent, Jonathan Shar, Trent Moorhead, Joseph Mundell, Imran Ghani
Mathematical and Computer Sciences Department
Indiana University of Pennsylvania, United States
E-mail : qhvx@iup.edu, zlhx@iup.edu, tcky@iup.edu, csxy@iup.edu, ighani@iup.edu
*Corresponding Author: Imran Ghani

Abstract

The earliest record of sports betting occurred over 2,000 years ago during the early years of the Greek Olympics. The main reason behind sports betting is simple - to make money. With the frequency of betting websites becoming more available and the legalization of online gambling, more and more people are betting to try their luck. When a user places a bet with a bookmaker, the bookmaker often uses certain techniques and shady odds to ensure that no matter what they will make money off the interaction. With new Artificial Intelligence and Machine Learning algorithms, the ability to create an Artificial Neural Network that helps predict the outcome of a sporting event is well within reach. The issue with programs widely available on the market today is they only pertain to one sport with a simple classification result, win or loss, with an accuracy rating around 55%. With those odds, the user making the bet would have close to the same odds if they were to guess. Having the option to choose between many different sports on one system and not limiting the software to either betting or predictions, allows for a new approach to both the betting and sports prediction software platforms. Combining the two allows for a much quicker and easier way to not only make decisions on which bets to make, but allows for a more accurate prediction in the same place without having to open two different software applications that each have a cost attached to them.

***Keywords*:** Machine Learning, Artificial Neural Network, Sports Betting, Artificial Intelligence

1. Introduction

Machine Learning (ML) algorithms use statistics to find patterns in big data. With sports betting becoming legal in some states, online vendors have jumped at the opportunity to provide an online environment for safe betting transactions. Using a dataset consisting of past games, players statistics, and game statistics; an accurate machine learning algorithm can be created. Using this ML algorithm[1], we can accurately predict the outcome of future games. Users will be able to pay a monthly subscription fee to access these results as well as find the best odds to bet on. This provides them with the highest chance of the highest profit.

2. Model-View-Controller Approach to Develop Sports Betting System

In this project, we implemented an MVC design pattern [1-3]. **Fig. 1.** presents the package diagram of our MVC design pattern with 3 packages in it.

2.1 Model Component

The model component holds files to give the website structure and better visuals for the users. CSS files allow for uniform formatting that can more easily be changed if needed.

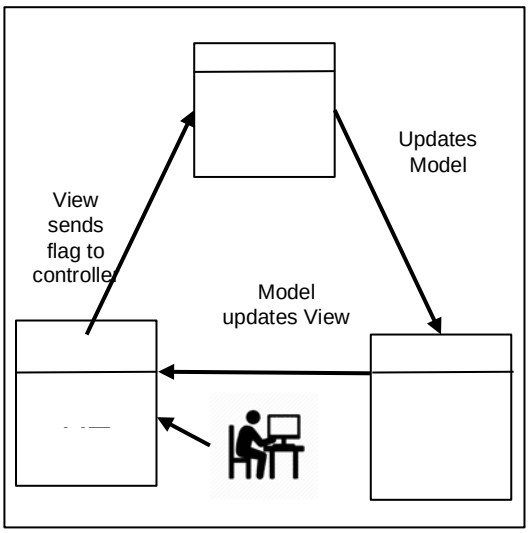


Fig. 1. MVC package diagram

2.2 View Component

The view component holds php files to incorporate the models and controllers together to create a comprehensive website. In the php files are also html programming to create more interactable components on the website.

2.2 Controller Component

The controller is a collection of "flags" used to direct the system to a certain file after the user clicks on part of the website. This allows for quick changes to any part of the website without having to change the entire system.

3. System User Interface

The system user interface will allow users to customize their User Interface for a more personal experience. The user will have the ability to customize their dashboard with a drag and drop interface where each module can be positioned to the users preference. Customization features include colors, positioning, shrinking, expanding, etc. Users also have the capability to create, add, or delete modules on their dashboard. In the dashboard users will be able to see their favorite teams, players, stats, current trends, etc.

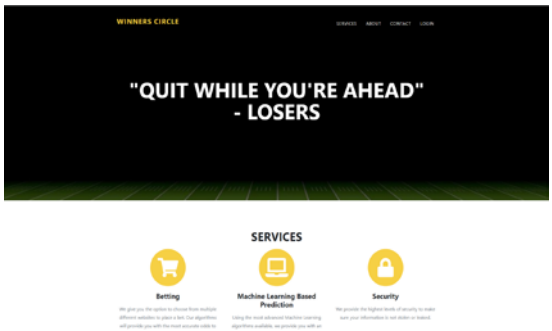


Fig. 2. GUI of Dashboard

In order to overlay the virtual 3D models on the AR system, **Fig. 2** a marker in the input image should be detected first.

4. Kanban-based Sprint Planning using Trello

We used a Kanban board using Trello software as shown in **Fig. 3.** We created three (3) sprints for this class project to get experience in SCRUM team programming.

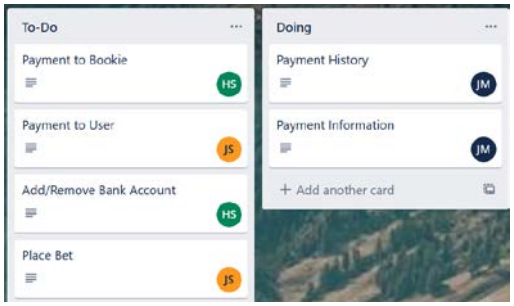


Fig. 3. Trello view of Tasks

5. Conclusions

In this paper, we present an interactive system that displays sports information through machine learning and sports betting by using the Model-View-Controller approach.

References

[1] Manuel Silverio, "My Findings on Using Machine Learning for Sports Betting" https://towardsdatascience.com/my-findings-on-using-machine-learning-for-sports-betting-do-bookmakers-always-win-6bc8684baa8c, 2020

[2] R. Bunker, T. Susnjak, "The Application of Machine Learning Techniques for Predicting Results in Team Sports" , https://arxiv.org/pdf/1912.11762.pdf , 2019

[3] Stübinger, J.; Mangold, B.; Knoll, J. Machine Learning in Football Betting: Prediction of Match Results Based on Player Characteristics. *Appl. Sci.* **2020**, *10*, 46. https://www.mdpi.com/2076-3417/10/1/46

MVC based Approach to Implement Online Educational Plant Content Management System

Neil Yoder, Brooklyn Kotula, Seth Edwards, Michaela Lutz, Imran Ghani
School of Math & Computer Sciences
Indiana University of Pennsylvania, United States
E-mail : xcdy@iup.edu, qytw@iup.edu, pgkbc@iup.edu, qpyx@iup.edu, ighani@iup.edu
*Corresponding author: Brooklyn Kotula

Abstract

Poisonous and harmful plants have coexisted with Humans for thousands of years, and with them comes a need to study and understand them. Anywhere you go in the world you will find some variation of vegetation that will cause irritation when touched. Existing educational systems can lightly cover this and help teach us what we should or should not touch. However, none of these systems go in depth enough to help people truly understand what is dangerous and what is not. We are striving to both educate our users with lasting knowledge that could one day help them and provide a bank of information that can be called upon when needed. We propose an education system, Rash Decision, that will offer a guide to users through an interactive component-based system, designed to help build lasting memory while also providing a source of fun entertainment. It will focus on quiz-based lessons that give you hands-on experience viewing different plants and learning about them as you go along. The system will challenge users to differentiate between poisonous and harmless vegetation and encourage them to try to top their previous high score. Developing this could lead to an increased knowledge of our environment and increased health among its users.

***Keywords*:** Education System; Harmful Plants; MCV based; Software Product Line

1. Introduction

Current and future generations are increasingly working with interactive learning technologies. These are rapidly becoming powerful multisensory, interactive environments that challenge current educational values[1]. Interactive environments are those at which users interact with digital content which drives a user experience which is engaging, memorable, informative and inspiring[3]. Interactive environments can help educate users on poisonous plants and build a long-lasting memory of information while also providing a foundation of entertainment. Interactive learning is a hands-on, real-world method to education. Interactive learning actively engages the students in wrestling with the material. It strengthens the classroom for both students and faculty. Lectures are changed into discussions, and students and teachers become partners in the journey of knowledge acquisition[2]. By creating an interactive quiz to test users on identifying poisonous plants, this could lead to an enhanced knowledge of our environment and improved health among our users while enjoying themselves.

2. Model-View-Controller Approach to Develop RashDecision System

In this project, we implemented MVC design pattern **Fig. 1.** presents the package diagram of our MVC design pattern with three packages in it.

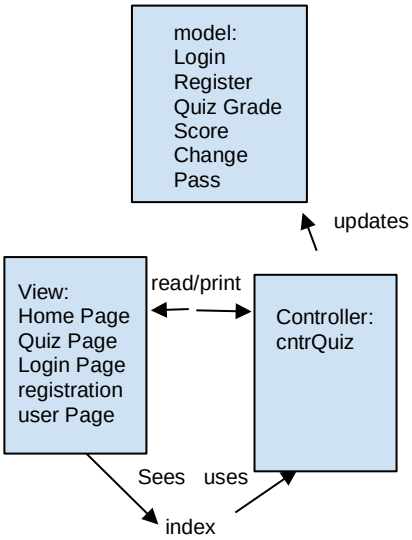


Fig. 1. MVC package diagram

2.1 Model Component

Our model package is designed to handle all data calculations and manipulations. It also includes all changes that are made to the external database.

2.2 View Component

Our view package displays all of the content in our system. It handles all of the external and internal pages and displays that to the user.

2.2 Controller Component

The Controller package deals with all traffic and links the other two packages together. It directs what page to go to and what model components to call when a calculation is needed.

3. System User Interface

This is the homepage that has links and tabs to other various options provided by RashDecision.

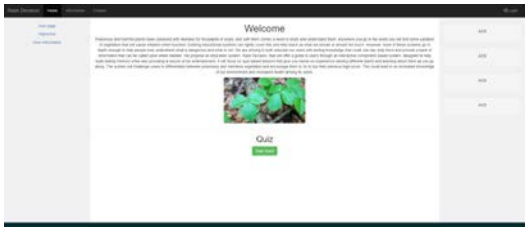


Fig. 2. GUI of HomePage

In order to access the full content of RashDecision, **Fig. 2** shows our current functionality and makes navigating RashDecision easier.

4. Kanban-based Sprint Planning using Trello

We organized our project with a Kanban board using Trello software as shown in **Fig. 3.** We created three (3) sprints for this class project, each with outlined objectives in order to keep the project on-schedule, as well as defined resulting product increments.

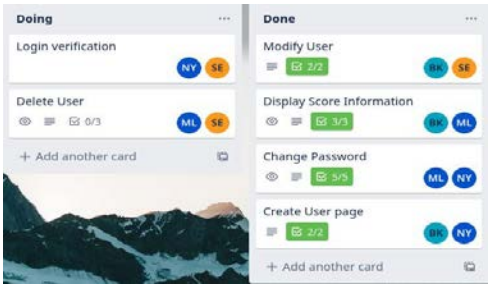


Fig. 3. Trello view of Tasks

5. Conclusions

In this paper, we present a method to educate and provide a system for educators to effectively display information through an interactive process.

References

[1] Laszlo, Alexander, and Kathia Castro. “Technology and Values: Interactive Learning Environments for Future Generations.” *Educational Technology*, vol. 35, no. 2, 1995, pp. 7–13. *JSTOR*, www.jstor.org/stable/44428957. Accessed 12 Nov. 2020.

[2] "What is Interactive Learning? - Overview & Tools." Study.com, 10 August 2016, study.com/academy/lesson/what-is-interactive-learning-overview-tools.html.

[3] Lusens. (2017, October 16). Interactive Environments: Interactive Spaces: Immersive Environments. Retrieved from https://lusens.com/interactive-environments/

KSII The 12th International Conference on Internet (ICONI) 2020.

MVC based Approach to Implement Web-based Food Ordering System

Preston Dulion, Luke Tvarozna, Joseph Gula, Nicholas Villano, Imran Ghani
School of Computing, Faculty of Engineering
Indiana University of Pennsylvania, USA
E-mail : cyky@iup.edu, rfrw@iup.edu, vmty@iup.edu, gxjy@iup.edu, ighani@iup.edu
*Corresponding author : Preston Dulion

Abstract

Online food delivery has been around since late 1995, which started with World Wide Waiter, created by Craig Cohen and Michael Adelberg at Stanford University. This programs goal is to make easy and accessible way to view the menu for multiple restaurants and to get them delivered to your front door all without needing to leave the app. In Food 'N' Now, we offer accurate delivery times, secure checkout, and a wide variety of both local restaurants and popular chain restaurants. Normally food delivery is handled by the restaurants themselves, most popular in pizza restaurants. Our app allows you to go around their delivery system and always gives the customer first priority. Food delivery services help prevent illnesses and diseases due to less contact. Our food delivery allows the customer to pick their time of delivery and how it will be presented either face to face or no contact delivery. Food 'N' Now application will be able to accurately track when the food is ready for delivery, as well as tracking the driver. Our login is secured with high level encryption, as well as mandatory two-factor authentication. Our customers will be able to securely checkout and we accept all major credit/debit cards, as well as PayPal. This paper proposes a new application-based food delivery service to be using the Model-View-Controller (MVC) architecture. Food 'N' Now is a web server-based software and will be used to make/get food orders, and as a means of contact between the customer and the driver, and the driver and the restaurant.

Keywords: Model-View-Controller

1. Introduction

Online food delivery is a conveince that hasn't been around for too long. The first instance being World Wide Waiter [1]. Online food delivery allows customers to order food from whereever they are, and schedule it to be delivered, without needing to talk to anyone. Coders' Igloo's Food 'N' Now system of online food delivery intends to follow the examples of other popular food delivery services, like DoorDash [2].

The intent of the Food 'N' Now project is to be a class project to better understand how websites are created. As well as be a learning experience for how Model-View-Controller [3] approaches should function.

2. Model-View-Controller Approach to Develop the Food 'N' Now System

In this project, we implemented MVC design pattern [3]. **Fig. 1.** presents the package diagram of our MVC design paterm with 3 pacakges in it.

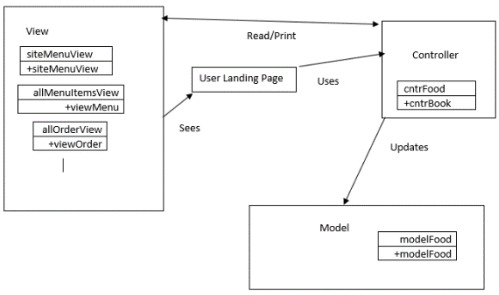


Fig. 1. MVC package diagram

2.1 Model Component

We use two models one for login which will welcome the user with their username and one for food showing all components of name, restaurant, description, and cost.

2.2 View Component

Our MVC uses three different view components. The first view is for the landing page when the site is accessed, the second view displays items from our available restaurants, and the third view shows the user their orders.

2.3 Controller Component

The controller is made to be a seamless process of logging in, viewing our restaurant menus, and ordering items.

3. System User Interface

Fig. 2. GUI of Dashboard

Fig. 2 shows the landing page a customer would receive after they have logged in. The user is able to order food, as well as just browse the menu. Clicking on the FOOD 'N' NOW brings them back to this page.

4. Kanban-based Sprint Planning using Trello

We used Kanban board using Trello software as shown in **Fig. 3.** We created three (3) sprints for this class project.

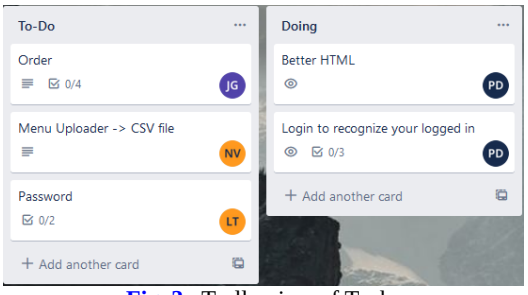


Fig. 3. Trello view of Tasks

5. Conclusions

In this paper, we present the website for our online food delivery service known as Food 'N' Now. The Food 'N' Now system is being developed through a Model View Controller approach. PHP and HTML are the two primary programming languages that are being utilized in the development of this website. The Food 'N' Now system intends to be competitive in the online food delivery market against major companies that dominate the market such as DoorDash, Uber Eats, Postmates, and Grubhub.

References

[1] Waiter.com, Inc. "Our Story Happiness Through Food" https://www.waiter.com/our-story, 1995-2020

[2] DoorDash Inc. "Delivering Good" https://www.doordash.com/about/, 2013-2020

[3] Davis, I. (2008, December 8). What Are The Benefits of MVC? Retrieved from Internet Alchemy: https://blog.iandavis.com/2008/12/what-are-the-benefits-of-mvc/

MVC Based Implementation for Online Surveys

Gayah Karmo, Jeffrey Cook, Joseph Frantzen, Zachary Boyer, Imran Ghani
Department of Computer Science
Indiana University of Pennsylvania, Pennsylvania
E-mail : gdfv@iup.edu, ltnw@iup.edu, nzky@iup.edu, cqcy@iup.edu, ighani@iup.edu
*Corresponding authors: Zachary Boyer

Abstract

Online Surveys have been around since the early 1990s, and since then they have been continuously used throughout a variety of fields for data gathering. Today the most popular version of online surveys are web-browser based and provide immediate feedback to the survey owners. One issue with modern online surveys is a lack of a great user experience. This includes but is not limited to, poor customer service, lack of accessibility, and poor customization. With Survey Central (SC) we plan to make a questionnaire generator that can be used by anyone from those with no technological knowledge to those in the computer science field. Depending on your knowledge and willingness to invest time you can create whatever you want. Along with this one large part of our company and product will be based on and around customer service, this includes a live chat, email, and phone support. Survey Central is a product based on designing, hosting, and producing questionnaires. The product will help format and design the questionnaires along with hosting the surveys and their answers. The software will be hosted on a website with a user interface and account system for all users. The benefits of hosting and creating surveys online is the ease of use for the survey takers and accessing the answers for the survey creators. By implementing these features in a well managed and responsive system we create an opportunity to provide a needed service to those looking for the best Survey Creation software available.

***Keywords*:** Online Surveys; Web-Browser Based; Survey Central; Questionnaire; Survey

1. Introduction

Computerized surveys have been around since about the 1970s where the first computer based surveys were taken via punch cards[1]. Online surveys is a powerful and easy tool to use for making surveys. The purpose of online surveys is for businesses and people to be able to get results based on what others think about a certain subject[3]. There is a great benefit for surveys in the medical field too. A group of surgeons implemented a brief cross-sectional self-reported questionnaire that addressed the areas of clinical examination and decision-making skills of management of equinus deformity in CP children". The data you can learn from a survey is very crucial and we can develop a survey software that people can use to save lives[4]. With this in mind we decided to create a software to allow users to create detailed and custom surveys to collect information.

2. Model-View-Controller Approach to Develop Survey Central System

In this project, we implemented MVC design pattern. **Fig. 1.** presents the package diagram of our MVC design pattern with 3 packages in it.

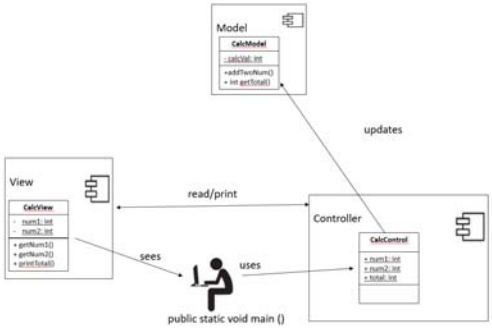


Fig. 1. MVC package diagram

2.1 Model Component

We designed and implemented our model component to manage and implement our control component in a neat and concise way. This provides some separation of our files and allows us to troubleshoot and debug easier knowing exactly how and where certain aspects of our program are handled[2].

2.2 View Component

For our view component we use basic JFrame structures to create a GUI for our users to log into and work from. This simple yet effective method allows us to use a language we are familiar with while providing just enough flexibility to allow our users to work the software efficiently.

2.3 Controller Component

Our controller component brings the entire software together, this does the basic tasks such as utilizing some view components and provides a place to put all methods that do raw processing apart from our model components.

3. System User Interface

Our system user interface uses a basic dashboard and login system. The login system provides an online aspect to the software to verify credentials and allow a user to store their data online via our database.

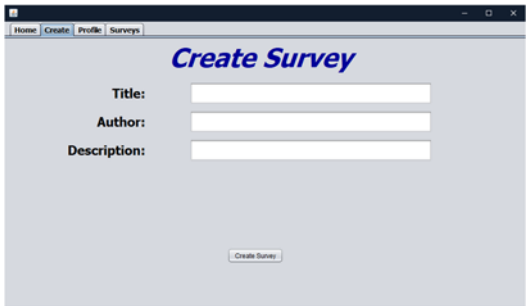


Fig. 2. GUI of Dashboard

Our GUI as discussed before is basic JFrame forms created via Netbeans, this allowed us to quickly create a detailed and well-made GUI for us to then implement our logic into.

4. Kanban-based Sprint Planning using Trello

We used Kanban board using Trello software as shown in **Fig. 3.** We created three (3) sprints for this class project and at the end of each sprint we analyzed and restructured our methods.

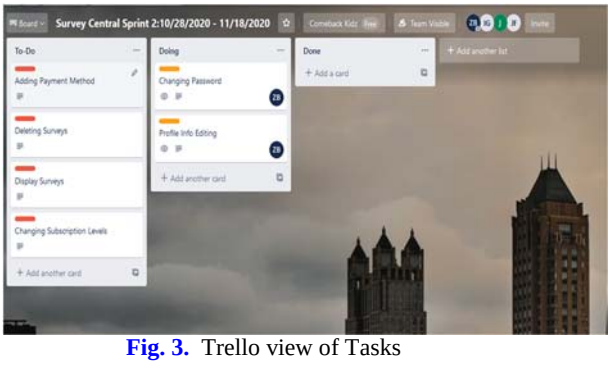


Fig. 3. Trello view of Tasks

5. Conclusions

In this paper, we presented an online and editable version of creating, distributing, and taking surveys and questionnaires for data collection and analysis. We hope to create a program capable of gathering as much data as possible across all fields of technology, medicine, construction, ecology, etc.

References

[1] Jakub Wierusz,"Online surveys - short history" https://www.surveylab.com, 2020.

[2] Steve Smith, "Overview of ASP.NET Core MVC" , https://docs.microsoft.com, 2020

[3] "Online Surveys: Definition, Characteristics, Examples, Advantages and Disadvantages" https://www.questionpro.com/blog/what-are-online-surveys/

[4] Gendy, S; ElGebeily, M; El-Sobky, TA; Khoshhal, KI; Jawadi, AH "Current practice and preferences to management of equinus in children with ambulatory cerebral palsy: A survey of orthopedic surgeons" https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6394235/

The impact of online business service quality of Chinese snack foods on consumer satisfaction and willingness to buy back

Dandan Yu and Lanmeng Wu
Department of Economics, Peking University
China
[e-mail: 332090197@qq.com, 332090197@qq.com]
*Corresponding author: Lanmeng Wu

Abstract

In recent years, the proportion of leisure food in E-commerce platforms has kept increasing year by year for its attribute characteristics and the rapid development of E-commerce platforms, but some problems have also emerged. This paper aims at studying the influence of service quality on customer satisfaction and repurchase intention by means of specific division of service quality of E-commerce with leisure food E-commerce being an example. In this paper, the collected data is analyzed in terms of reliability and validity, factor analysis and construction of equation models in the form of questionnaires with software SPSS and AMOS. The conclusions obtained from analysis include: 1. Customers' expectation before purchasing has a significant positive impact on consumer satisfaction. 2. The assumption that E-commerce web design has a significant positive impact on consumer satisfaction is false. 3. E-commerce logistics has a significant positive impact on consumer satisfaction. 4. The quality of e-commerce products has a significant positive impact on consumers. 5. E-commerce after-sales service has a significant positive impact on consumer satisfaction. 6. Consumer satisfaction has a significant positive impact on repurchase intention.

***Keywords*:** Leisure Food E-Commerce, Influence Factors, Satisfaction Degree, Repurchase Intention

1. Introduction

The worldwide fast food market is growing rapidly in recent years. 2010-2014 witnessed the rapid development of leisure food industry of China, as the CAGR is up to 12.7% and the market size in 2014 is getting close to 100 billion. Despite the slow-down growth of the industry in the previous years, it has maintained a growth rate of more than 6%, and the current market size is up to 109.5 billion Yuan. The huge market size creates a promising prospect for the development of leisure food. For the disadvantage of traditional offline sales of leisure food restricted by factors such as region and sales channels, the development of E-commerce provides the industry of leisure food with a significant chance for development.

2. Theoretical Background and Research Model

2.1 Theoretical Background

Consumer satisfaction degree is considered as a feeling gained from evaluating the products and the consumption experience. [1] It refers to the difference between consumer expectation/belief and the product performance. It is affected by

enterprise image, service, product quality, price, etc. According to the literature review, this paper agrees the opinion proposed by Oliver in 1997 that consumer satisfaction degree reflects the psychological state when the consumer's expectation and demand are satisfied. It is a judgment by the consumer that how well the product or the service has satisfied their needs. [2] Repurchase intention refers to the willingness of the consumer to maintain transaction relationship with the enterprise, that is, after the consumer purchases and uses the product, he would evaluate the shopping and user experience to judge whether he has the intention to purchase the product of the company again.[3] Consumer satisfaction degree is considered to positively affect consumer's repurchase behavior.

2.2 Research Model

Based on the collation and summary of the prior research, this research, in consideration of the actual process of consumers' purchase behavior in E-commerce, holds that customer expectation, web design, logistics, product quality and after-sales service are five major factors with influence on consumer satisfaction and consumer satisfaction has a significant impact on repurchase intention. The research model of this paper is as shown in the **Fig. 1**.

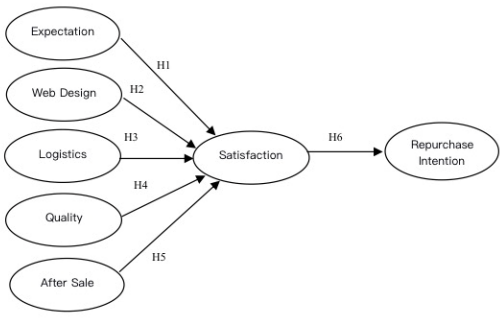


Fig. 1. Research Model.

H1: Consumer expectation has a significant positive impact on consumer satisfaction.
H2: E-commerce web design has a significant positive impact on consumer satisfaction.
H3: E-commerce logistics has a significant positive impact on consumer satisfaction.
H4: The quality of E-commerce products has a significant positive impact on consumers.
H5: After sale has noticeable positive influence on consumer satisfaction degree.
H6: Consumer satisfaction has a significant positive impact on repurchase intention.

3. Research Approach and Results

3.1 Research Approach

The author conducted a questionnaire survey on customers with purchase behaviors of leisure food on E-commerce platforms of China in terms of the subject from September 2, 2019 to September 27, 2019. Among the 278 questionnaires distributed in the survey, 270 real valid questionnaires were recovered after data examination at a recovery rate of 97%. Reliability analysis, factor analysis and construction of equation model analysis are conducted with SPSS22.0 for survey data.

References

[1] Hunt and H. Keith, "CS/D-overview and future research directions," Conceptualization *and Measurement of Consumer Satisfaction and Dissatisfaction*, Vol 5, No. 16, pp. 456-487, 1977.
[2] Richard L. Oliver, "Satisfaction: a behavioral perspective on the consumer," New York: McGraw-Hill, 1997.
[3] C. Mingliang, "An Empirical Study on the Determinants of Customers' Repeated Purchase Intentions at Different Stages of the Life Cycle," *Management world*, Vol 11, No. 12, pp. 93-99, 2002

Research on Satisfaction Evaluation Based on Tourist Big Data

Rui LIANG, Hanwen GUO, Zeyu JIAO, Ziyang LIU
Business School Beijing Union University, Beijing, China.
[e-mail: rui.liang@buu.edu.cn, 921106GHW@gmail.com, 447052870@qq.com, morninglzy@hotmail.com]
*Corresponding author: Ziyang LIU

Abstract

With the improvement of people's living standards and the development of tourism, tourists have greater freedom in choosing destinations. Therefore, as an indicator of satisfaction with scenic spots, tourist comments are becoming increasingly prominent. In this paper, the Five Great Mountains in China are selected as analysis objects, which refer to the Mount Hua in Shaanxi, the Mount Thai in Shandong, the Mount Song in Henan, the Mount Heng in Hunan and the Mount Heng in Shanxi. The online reviews of tourists on the Online Travel Agency (OTA) website about the Five Great Mountains from 2015 to 2018 are collected as research samples. The text analysis method and R language are used to analyze the content of the tourist reviews, while the high-frequency words in the word cloud are used for visual display. In addition, the entropy weight method is used to determine the index weight and tourist satisfaction is evaluated to understand the weaknesses of those scenic spots. The results of the study show that firstly, the tourist satisfaction with the Five Great Mountains is basically consistent with its popularity. Secondly, through weight analysis, tourists pay special attention to the landscape features and environmental health of the scenic area, so that relevant departments should focus on building the landscape characteristics and improving the environmental health of the scenic area. At the same time, the accommodation and service management of the scenic spot cannot be ignored. Finally, according to the analysis results, suggestions are made on how to improve the tourist satisfaction with the Five Great Mountains.

***Keywords*:** Online Travel Agency, text analysis, word cloud, the Five Great Mountains, tourist satisfaction

1. Introduction

In recent years, China's tourism informatization construction has achieved considerable development. With the tourism online service market gaining scale, the marketing pattern with full-coverage tourism propaganda has basically

formed. Meanwhile, tourism e-government has been steadily advancing and tourism informatization has significantly strengthened the tourism industry [1]. Under the background that the combination of information technology and tourism turned into a new development hotspot, the smart tourism service system has become one of the focus of enterprise construction [2]. According to the China Scenic Tourism Research Report, at least 81.5% of tourists often share their travel experiences on WeChat, Weibo, OTA (Online Travel Agency), community forums, etc. and 84.8% of tourists share travel experiences instantly through social tools. Among tourists who often share travel experiences, 60.4% of them refer to the travel guide and arrange their travel itinerary according to the guide. As online travel has dominated China ’s tourism market, various OTA sites make an effort to take a slice of this market. In this condition, tourists use these OTA sites to evaluate the experience of scenic spots, especially after the 4G era spreading across different provinces in China. To a great extent, real-time reviews of scenic spots and the follow-up reviews both reflect the tourists' image perception and satisfaction about the scenic spot. In recent years, a large number of scholars have used online reviews to study customer satisfaction. Radojevic and other scholars have statistically analyzed the text data of hotel reviews and found that tourists are more concerned about the star rating, room layout and air conditioning equipment when they stay at the hotel[3]. Scholar Zhao analyzed tourist satisfaction with scenic spots in world cultural heritage sites such as the ancient city of Pingyao and Yungang Grottoes by using online reviews as a sample [4]. Through online data, Li took the key scenic spots in Xi'an as an example to construct a tourist satisfaction model and analyzed the tourists' satisfaction with the scenic spots in Xi'an [5]. From the previous research, there are more and more scholars who use network text data as a method of satisfaction research.

American scholar Cardozo first introduced the concept of "satisfaction" in marketing in 1965 to show that customer satisfaction will drive customer behavior [6]. After 1970, Pizam and other scholars applied the concept of satisfaction to the tourism industry, which provided a theoretical basis for future scholars to study tourist sites and tourist satisfaction [7]. Chinese scholars' research on tourists and satisfaction began in the late 1980s. Since the development of the tourism market should focus on tourists, many scholars put effort on investigating the way of improving tourists' satisfaction with scenic spots. The research on tourist satisfaction in this paper mainly uses the definition put forward by Pizam and other scholars, which shows that tourist satisfaction is the comparison result of tourists 'expectations and actual experience. If the travel destination almost meets the tourists' expectations, it indicates that the tourists are satisfied. Otherwise, it means tourists are not satisfied [8].

The Five Great Mountains in China has always been a popular scenic spot for tourists. Therefore, scholars in China have never stopped doing research on this famous attraction. The research content mainly includes tourist satisfaction with scenic spots, protection of scenic spot resources and tourism development, as well as cultural aspects of scenic spots. Most scholars use the method of qualitative research and some scholars conduct research in quantitative terms. However, quantitative analysis mostly focuses on the

tourist and relatively few studies focus on scenic spots [9] [10]. Based on the perspective of tourists, this article compares and analyzes the satisfaction of tourists in the five scenic spots of the Mount Hua in Shaanxi, the Mount Thai in Shandong, the Mount Song in Henan, the Mount Heng in Hunan and the Mount Heng in Shanxi. In the context of big data, it researches the needs of tourists based on the word-of-mouth from the third-party Internet. By monitoring OTA tourist reviews through the entire network, network reviews and public opinion data are deeply mined and data evaluation models are built to understand tourists' needs and emotional changes in a more scientific and intelligent way. In this way, the shortcomings of the development of these five scenic spots are found out and corresponding solutions are proposed, which helps marketers to carry out precise marketing to tourists. In addition, it could be beneficial to enhancing the image of the scenic area and providing a reference for the subsequent development of the scenic area [11] [12].

2. Research Method

2.1 Data Sources

The research data of this paper is mainly taken from the evaluation text of the scenic spots provided by the tourists who used the three platforms of Ctrip, Meituan and Qunar to buy tickets of the Five Great Mountains between 2015 and 2018. The paper uses Python software for data crawling and the text data was published from January 1, 2015 to December 31, 2018. The total amount of data initially obtained was 81254. Among the data, the number of data from the Mount Thai and the Mount Hua is 17641 and 17515 respectively, while the data number from the Mount Heng in Hunan is 15321 and the Mount Heng in Shanxi is 14229. The amount of data from the Mount Song is 16848. Then the data was selected by removing duplicate data and default praise. The principle of comment elimination is to clean reviews of irrelevant subject content, repeated reviews, reviews that are too radical or highly commercial and reviews that are irrelevant to visitor satisfaction, etc. Ultimately, the initial data for controlling subsequent calculation of satisfaction with the Five Great Mountains are all eliminated to 12,000, with a total data volume of 60,000. Compared with traditional methods such as questionnaire surveys, the advantages of the network text analysis method include the large sample size, rich content and fast acquisition speed. As a result, the text content is time-sensitive and more authentic. Whereas, there are some disadvantages including the narrow age range of the sample group and the inability to set subjective questions to investigate tourist satisfaction [13].

2.2 Research process

Based on the review sample, this paper mainly adopts the text analysis method according to the steps of sample data collection → word frequency statistics → project construction → satisfaction calculation → result analysis. The Python is firstly used to collect the review samples of the Five Great Mountains separately, then the invalid content of data is removed after the collection [14]. On the basis of data removal rules mentioned above, the data is imported into the R language software to use the word cloud package to perform text word frequency statistics (**Table 1**). In the final, the word cloud is generated for visual presentation (**Fig. 1**).

Fig. 1. High-frequency word cloud of the Five Great Mountains

landscape features, cultural characteristics, infrastructure and service management. According to the elimination criteria above, the remaining texts are all related to tourist satisfaction and emotional expressions. The text contained in the nine classified indicators is imported into R language and sentiment analysis is performed. The sentiment classification is divided into 3 comments, which are positive comments, neutral comments and negative comments. In this paper, the fuzzy comprehensive evaluation method and the normalization method are used to process the sentences of the sentiment classification. The weights of the three reviews are 1, 0.5, and 0.1. After calculation, the initial data of tourist satisfaction in the Five Great Mountains areobtained on removing the objective content in the comments. Finally, only one subjective content or topic is retained for analysis. The semantics are divided into transportation, catering, accommodation, shopping, environmental health, After drawing the word cloud, the text is semantically classified separately, which is based After calculating the initial data of visitor satisfaction, this paper uses index weights to reflect the different roles played by each evaluation index in the overall review. At present, the methods for determining weights are mainly divided into two categories. One is the qualitative weighting method, which includes the fuzzy comprehensive evaluation method, analytic hierarchy process, etc. The other is the quantitative weighting method, including entropy weight method, principal component analysis method, regression analysis and so on. The entropy weight method is mainly used in this paper to determine the index weight, which is a method to determine the objective weight according to the index variability. The entropy weight method has strong objectivity and is fully applied to various types of original data where weights need to be determined. The evaluation standard of the entropy weight method is that if the entropy value of a certain evaluation index is smaller, it means that the index change value is larger. The more information is provided, the greater effect the evaluation has and the greater the weight is. When the entropy value is large, the information it provides is reversed [15]. The process of calculating weights by entropy weight method is as follows.

Table 1. High-frequency characteristic words and word frequency statistics of the Five Great Mountains

Serial number	Item	Word frequency	Serial number	Item	Word frequency	Serial number	Item	Word frequency
1	泰山(Mount Thai)	8096	16	晚上(Night)	4556	31	体力(physical strength)	3667
2	日出(Sunrise)	7199	17	索道(Ropeway)	4458	32	便宜（Cheap）	3622
3	华山(Mount Hua)	6788	18	风光(Scenery)	4416	33	到达(Arrival)	3619
4	嵩山(Mount Song)	6542	19	景区(Scenic area)	4330	34	价格(Price)	3512
5	衡山(Mount Heng in Hunan)	6511	20	排队（Queuing）	4309	35	很美(Very beautiful)	3501
6	恒山(Mount Heng in Shanxi)	6231	21	景色（Sight）	4308	36	夜爬(Night climb)	3484
7	壮丽 (Magnificent)	5892	22	建议（Suggestions）	4282	37	天气(Weather)	3473
8	绝壁(Cliff)	5819	23	登山(Mountain climbing)	4200	38	步行(Walking)	3352
9	五岳(The Five Great Mountains)	5644	24	文化(Culture)	4143	39	网上(Online)	3330
10	爬山(Climb)	5398	25	俊美(Pretty)	3941	40	休息(Rest)	3299
11	风景 (Landscape)	5089	26	气势（Momentum）	3915	41	很累(Very tired)	3259
12	险峻(Steep)	4836	27	感觉（Feeling）	3877	42	景点 (Attractions)	3246
13	山上（On the mountain）	4734	28	缆车（Cable car）	3807	43	大巴(Bus)	3231
14	值得（Worth）	4725	29	时间（Time）	3783	44	取票(Taking tickets)	3222
15	特色（Features）	4687	30	门票(Tickets)	3749	45	攻略(Guide)	3198

(1) Standardize the initial values of 9 evaluation indicators of the Five Great Mountains.

Let αij (i = 1,2, ..., m; j = 1,2, ..., n) be the satisfaction value of the j-th index in the i-th evaluation object. The paper has 5 evaluation objects (The value range of m is 1-5) and the initial tourist satisfaction matrix A composed of 9 evaluation indicators (the value range of n is 1-9), then A is:

$$A = \begin{pmatrix} \alpha_{11} & L & \alpha_{m1} \\ M & O & M \\ \alpha_{1n} & L & \alpha_{mn} \end{pmatrix}$$

Normalizing the values in the A matrix to obtain α ', then a new matrix A' is generated.

$$A' = \begin{pmatrix} \alpha'_{11} & L & \alpha'_{m1} \\ M & O & M \\ \alpha'_{1n} & L & \alpha'_{mn} \end{pmatrix}$$

Among them, the standardized formula is as follows.

$$\alpha'_{ij} = \alpha_{ij} - \min \alpha_{ij} \Big/ \max \alpha_{ij} - \min \alpha_{ij}$$

$$\alpha'_{ij} = \max \alpha_{ij} - \alpha_{ij} \Big/ \max \alpha_{ij} - \min \alpha_{ij}$$

(2) Calculate the information entropy of each indicator.

$$H_j = -k \sum_{i=1}^{m} p_{ij} \ln p_{ij}$$

In the formula, $k = 1 \big/ \ln m$, $p_{ij} = \alpha' \Big/ \sum_{i=1}^{m} \alpha'_{ij}$

Table 2. Evaluation Index System of Tourist Satisfaction in the Five Great Mountains

Index	Mount Thai	Mount Hua	Mount Heng in Hunan	Mount Heng in Shanxi	Mount Song
Transportation	1	4	5	2	3
	1.00000	0.38739	0.00000	0.77187	0.54180
Catering	1	4	5	2	3
	1.00000	0.38248	0.00000	0.74352	0.56444
Accommodation	1	3	5	4	2
	1.00000	0.48366	0.00000	0.12049	0.63383
Shopping	1	5	4	3	2
	1.00000	0.00000	0.53104	0.63383	0.66167
Environmental health	5	2	1	3	4
	0.00000	0.60234	1.00000	0.33759	0.08028
Landscape features	2	3	4	5	1
	0.50278	0.12286	0.08091	0.00000	1.00000
Cultural characteristics	4	3	5	2	1
	0.51895	0.70734	0.00000	0.75689	1.00000
Infrastructure	1	4	5	3	2
	1.00000	0.41422	0.00000	0.57399	0.69844
Service management	5	1	3	2	4
	0.0000	1.00000	0.56826	0.85221	0.27447
Total tourist satisfaction	1	3	4	5	2
	0.61414	0.44863	0.43917	0.26604	0.60629

(3) Determine the weight Wj of each indicator.

$$W_j = \frac{1 - H_j}{\sum_{i=1}^{n} (1 - H_j)}$$

According to the above calculation results, the weights of 9 indicators (**Table 2**) and the satisfaction scores of the 9 indicators of the Five Great Mountains (**Table 3**) can be obtained. This paper then uses the weighted sum method to calculate the total satisfaction of tourists. The formula is as follows.

$$E_i = \sum_{i=1}^{9} \alpha'_{ij} * W_j$$

The larger the Ei value is, the higher the tourist satisfaction is. The calculation result is shown in **Table 3**. Under a certain item, the closer the value is to 1, the higher the tourist satisfaction of the indicator is, while the closer the value is to 0, the lower the tourist satisfaction is.

Table 3. Tourist Satisfaction Ranking and Calculation Results of the Five Great Mountains

Content	Index	Evaluation weight
Total tourist satisfaction in the Five Great Mountains	Transportation（0.08999）	Positive（1.0）
	Catering（0.08947）	
	Accommodation（0.13081）	
	Shopping（0.08035）	Neutral（0.5）
	Environmental health（0.15062）	
	Landscape features（0.19161）	
	Cultural characteristics（0.07940）	Negative（0.1）
	Infrastructure（0.08697）	
	Service management（0.10075）	

Data source: Author

3. Tourist satisfaction analysis

3.1 Satisfaction analysis of single scenic spot

Based on the data analysis results of Chapter 2 (ie, **Table 3**), the satisfaction of the Five Great Mountains is ranked. The Mount Thai, which is the "First of the Five Great Mountains and the First Mountain in the World", has obtained the highest ranking of multiple indicators and total tourist satisfaction. However, the scores of the two indicators of environmental health and service management are at the bottom, such as "vacation", "congestion", "high price", and "extra charge". Some scenic spots have irregular behaviors of extra charges after entering the mountain. The management of the surrounding taxis is not perfect and there is a phenomenon of slaughtering customers. This shows that the environmental health problems caused by the flow of people and the difficulty in providing quality services in scenic spots still exist. As a result, it is important for the Mount Thai to increase human and material resources to improve the quality of service management in scenic spots and ensure that the environmental health on the mountain is good.

The Mount Song, which ranks second in overall satisfaction, is an important birthplace of Chinese civilization. As the World Geopark, it has undoubtedly been praised by the vast number of travelers in terms of landscape features and cultural characteristics, which are both in the first position of the index. From text analysis, this is mainly due to the combination of the Mount Song's own strong cultural atmosphere and the magnificent mountain scenery. In addition, the Shaolin martial arts bring a more profound tourism experience to tourists. In terms of other

indicators, the Mount Song ranks in the middle. Like the Mount Thai, it has certain deficiencies in environmental health and service management. For example, "fewer toilets", "rough rest areas" and "fewer garbage bins" need to be further improved.

The Mount Hua, ranking the third in total satisfaction, is the most challenging mountain of Chinese civilization. Data analysis of the collected text shows that the Mount Hua is in the middle of many indicators, but it has won the first place in service management. However, in the shopping experience, it is at the end of the ranking. According to the content of reviews, when tourists lack food and water supplement in high-altitude areas, some tourists encounter the situation that the price of goods is raised at the transaction place, which will inevitably cause great dissatisfaction among tourists. The management department should pay attention to the price management of shopping in scenic spots. Under the premise of fully considering the difficulty of transportation of goods at high altitudes, it should appropriately grant certain subsidies and formulate uniform price standards. What is more, it should implement price supervision and dynamic inspections. If they find that the price is raised at the transaction place, the legitimate rights and interests of tourists should be protected timely.

As a national nature reserve, the Mount Heng in Hunan ranks fourth in total satisfaction and ranks first in the Five Great Mountains in terms of environmental health, indicating that the scenic area attaches great importance to management and invests a lot of manpower and material resources. However, the satisfaction in transportation, catering and accommodation is not high. The complaints are mainly reflected in the rapid speed of mountain sightseeing cars and the rugged mountain roads causing tourists to experience motion sickness, etc. Some tourists have proposed that the bus lanes lack guardrails. Therefore, to a certain extent, it causes panic for tourists. In terms of catering, the poor taste of the meals served in the affiliated restaurants of the scenic area does not reflect the characteristics of Hunan cuisine. With regard to accommodation, the evaluation shows that the environment is shabby and expensive. In addition, the price of burning incense in newly-developed temples is too high, so it is necessary for the scenic spot to strengthen the management.

As the only 4A-level scenic spot in the Five Great Mountains, the Mount Heng in Shanxi is the last in the ranking of total tourist satisfaction. The main problem is that it has a relatively low development level. For example, the road to the mountaintop is not long and the time to climb the mountain is short. Tourists who have visited the other four mountains do not rate the Mount Heng in Shanxi very high caused by the large difference. In addition, the Mount Heng in Shanxi is located in Hunyuan county, Datong city, Shanxi province. The long round-trip drive also leads to dissatisfaction of some tourists. However, in addition to the landscape features of the Mount Heng in Shanxi, the ranking of other indicators is in the middle, indicating that the local scenic spots are well managed and there is something worth learning.

3.2 Analysis of factors influencing tourist satisfaction

In the index system for tourist satisfaction evaluation of the Five Great Mountains (**Table 2**), it can be seen from the analysis of index weights that the accommodation, environmental health,

landscape features and service management of the scenic spot reached 57% of the total. Among them, the biggest impact on tourists is landscape features (0.19161), followed by environmental health (0.15062) and accommodation (0.13081). Service management (0.10075) and transportation (0.08999) rank fourth and fifth respectively, which also have a great impact on tourist satisfaction with a scenic spot. With the gradual improvement of people's living standards, their requirements for the quality of scenic spots are also constantly improving. As a result, it is important to enhance these five aspects which are beneficial to tourists' perceived quality.

4. Recommendations and suggestions

Based on the above analysis, this paper proposes the following recommendations and suggestions.

1. Strengthen the information construction of scenic spots. The Internet has brought about historic changes and smart tourism has opened up new ways for scenic spots to go up in service and management. Therefore, in order to improve the competitiveness of scenic spots, it is important to increase investment in information construction and improve the intelligent service. As a result, big data collection and analysis should be fully used to regularly collect and analyze feedback information from tourists. In this way, the problems in scenic spots can be rectified specifically by paying attention to the opinions and suggestions of tourists. For example, improve the software and hardware in the service, improve the infrastructure and transportation facilities, strengthen the traffic guidance for tourists in peak seasons, strengthen the health supervision and inspection, create a safe, comfortable and orderly tourism environment for tourists [16], comprehensively improve the satisfaction of tourists, focus on management benefits and realize the win-win situation of the economic and social benefits of scenic spots [17].

2. Strengthen the construction of the tourism talent team. Tourism service is a window for a city to demonstrate its civilization, while tourism practitioners are the most direct propagators and bearers of a city's civilization. Therefore, tourism functional departments and scenic area management should attach great importance to training and continuously improving the professional skills and professional ethics of tourism practitioners. They should be aimed at new problems, new situations and new requirements in service. In this situation, the training work should be classified, planned and carried out in a targeted way to indeed improve the actual effect of the training. Training cannot be accomplished overnight and the improvement of service quality and tourist satisfaction is always on the way [18].

3. Highlight the characteristics and broaden the channels of tourism projects. The scenic spots in the Five Great Mountains should actively develop new tourism products according to their characteristics and increase the added value of tourism projects. Highlighting features is an effective way to attract tourists and increase the added value of tourism. There are three ways of emphasizing features of scenic spots. The first is to develop dishes with local characteristics to satisfy the tourists ’experience and pursuit of local cuisine. At the same time, strengthen the supervision and inspection of the prices in the scenic spots to protect the legitimate rights and interests of consumers. The second is to create local-style homestays and seek to combine with

poverty alleviation projects effectively. For example, the local government can give certain policy and financial support to formulate homestay service standards, strengthen training homestay service personnel, strengthen health supervision and inspection and use self-media or other tourist-like ways to guide tourists to experience, so as to achieve a win-win situation by combining poverty alleviation and characteristic homestay project. Thirdly, developing tourist souvenirs with unique local characteristics and cultural characteristics and avoiding similarities in the development of new tourism products. In this way, special products in the Five Great Mountains can be created, so as to enrich the tourist souvenir market and satisfy tourists' shopping demand, as well as protecting intellectual property rights.

4. Promote the culture of the Chinese nation and tell stories of the Five Great Mountains. The Five Great Mountains are the carrier of inheriting and carrying forward the Chinese culture. Every mountain with a long history and culture is an important part of Chinese civilization. Its spirit and cultural connotation have been circulating for thousands of years. Therefore, telling a good story of the Five Great Mountains, inheriting the essence of the millennium culture and advancing with the times to give the new era connotation are not only the responsibility but also the obligation of every tourist practitioner in the scenic spot. Every tourist practitioner in the scenic spot must fulfill its obligations with good professional qualities and professional ethics, in order to become a powerful promoter of improving the cultural self-confidence of the Chinese nation.

5. Improve the reception capacity of tourists in the peak period of scenic spots. Most tourists will choose to go to the scenic area for a holiday. In order to deal with the huge passenger flow, in addition to receiving tourists properly in the reception area of the scenic spots, they must also prepare for emergency situations and safety accidents. In the statistical text, some tourists complained that there are too many people in the holiday, which is too crowded and difficult to walk. There are also comments that poor security facilities could not guarantee the safety of a large number of tourists. This has more or less reduced tourists' satisfaction with the scenic area. Therefore, scenic spots should make good use of modern technology. For example, they can analyze the tourist capacity and environmental carrying capacity of scenic spots to ensure the safety and order of scenic spots. In addition, it is important to reasonably receive tourists and improve the regulatory level and dynamic prediction ability of scenic spots, so as to guide the flow of tourists scientifically.

5. Conclusions and Limitations

Based on the content of the review, this paper uses OTA review data to build a tourist satisfaction evaluation index system to evaluate the tourist satisfaction of the Five Great Mountains. The R language and other software are also used to conduct a text analysis of the tourist reviews. Through comparative analysis, we got the following conclusions. (1) Through horizontal comparison, it can be found that the ranking of tourist satisfaction among the Five Great Mountains in order is from the Mount Thai, the Mount Song, the Mount Hua, the Mount Heng in Hunan to the Mount Heng in Shanxi. In addition to the factors of the scenic spot itself, tourist satisfaction is also affected by local policies and geographical locations. (2) Based on

the factors, the analysis text shows that the focus of tourism is on the index of landscape characteristics, followed by great attention to the environmental protection of scenic spots. However, the data sample is small in the aspects of catering and shopping in the scenic area. According to the classification results of emotions, these two indicators have low scores with mostly negative emotions. In addition, the local government should strengthen the "four constructions and one promotion" of the Five Great Mountains, which is strengthening cultural characteristic construction, strengthening talent team construction, strengthening infrastructure construction, strengthening informatization construction and promoting the scenic area 's smart tourism [19]. It is also important to consolidate the leading position of the Five Great Mountains in the province and vigorously promote the development of the tourism economy of the province, so as to achieve the purpose of promoting the economic development of the province.

There are still some limitations in this study. Firstly, in terms of data samples, the data collected from three travel platforms needs to be further expanded. Since the user information of tourists cannot be identified, it is difficult to determine the representativeness of the samples, so that it is impossible to determine the satisfaction of tourists from a scientific perspective. Secondly, the paper does not categorize the content of the reviews in time, but only calculates the content and semantics according to the weight score during the calculation. For example, some tourists made suggestions in 2015 and the scenic spot management department may have rectified the problem in 2016, so that some scenic spot projects cannot obtain an absolute objective score when doing the calculation, which affects the tourist satisfaction ranking of the scenic spot. Finally, with regard to research methods, this paper mainly uses text analysis, followed by semantic classification and sentiment analysis to screen the text. However, this method can only passively obtain the satisfaction and dissatisfaction of tourists expressed in scenic spots. In the future, we can conduct on-site investigations and social surveys on key low-score project indicators after data analysis. In this way, we can study how to improve low-score projects to contribute to enhancing the satisfaction of scenic spots.

References

[1] Gov.cn., "*Boosting Tourism Development With Informatization Construction-Interpretation Of The "Thirteenth Five-Year Plan For National Tourism Informatization,"* 2017. [online] Available at: <http://www.gov.cn/> [Accessed 6 May 2020].

[2] Zuo, D.J., "The 13th Five-Year Plan for National Tourism Informationization," *China Tourism News*, 2017.

[3] Radojevic T. and Stansic N., "Ensuring positive feedback: factors that influence customer satisfaction in the contemporary hospitality industry," *Tourism management.* Vol 51, No. 1, pp. 13-21, 2015.

[4] Zhao P.Y., Cui Q., Zhang H. et al., "Research on Tourist Satisfaction of Shanxi World Heritage Tourist Destinations Based on Internet Reviews," *Shanxi Normal University* (Natural Science Edition), Vol 30, No. 3, pp. 64-69, 2016.

[5] Li L.M., Wang X.F., Liu Y. et al., "Model construction of tourist satisfaction evaluation in key scenic spots in Xi'an," *Value engineering*. Vol 31, No. 25, pp. 285-287, 2012.

[6] Cardozo R. N., "An experimental study of customer effort expectation and satisfaction," *Marketing Research*, Vol 2, No. 3, pp. 244-249, 1965.

[7] Pizam A., "Tourism 's impacts: the social costs to the destination community as perceived by its residents," *Travel Research*, Vol 16, No. 4, pp. 8-12, 1978.

[8] Pizam A., Neumann Y., Reichel A., "Dimensions of tourist satisfaction with a destination area," *Annals of Tourism Research*, Vol 5, No. 3, pp. 314-322, 1978.

[9] Shen Y. and Su Q., "Research on the satisfaction evaluation of tourist items in ethnic villages: Taking Longji Scenic Spot in Longsheng County, Guilin, Guangxi as an example," *Chongqing University of Science and Technology* (Social Science Edition), Vol 5, No. 1, pp. 101-102, 2009.

[10] Liu N. and Wu Z.J., "A study on the current situation of tourism satisfaction of Longji terraced scenic spots in Guilin, Guangxi," *Dazhong Science and Technology*, Vol 17, No. 5, pp. 162-164, 2015.

[11] Xu Y.H., You X.B., Wang Y.N., "Research on Tourist Satisfaction Evaluation Based on ROSTCM Method-Taking 5 Domestic Terraced Fields as an Example." *Tourism Forum*, Vol 11, No. 5, pp. 23-24, 2018.

[12] Gao J., "Research on marketing strategy based on Internet word-of-mouth in the context of big data-taking Yangzhou Slender West Lake as an example," *Jiangsu Business Review*, Vol 8, No. 1, pp. 34-41, 2018.

[13] He Y. and Yang X.X., "Application and comment of content analysis method in tourism research in China," *Southwest Yongye University*, Vol 10, No. 4, pp. 6-12, 2012.

[14] Sun Y., *"A Study on the Satisfaction of Outbound Tourists Based on Data Mining on the Internet Social Platform,"* Shanghai: Shanghai Normal University, 2015.

[15] Huo D.L., Liu L.F., "Research on weights of early-warning indicators of social risks in Beijing Winter Olympics based on entropy weight method," *Shenyang Institute of Physical Education*, Vol 38, No. 5, pp. 47-55, 2019.

[16] Chen X.X., "Thoughts on improving the quality of tourism industry," *Shandong Academy of Agricultural Management Cadres*, Vol 2, No. 1, pp. 84-85, 2006.

[17] Zhu L., "Research on ticket price management of public scenic spots from the perspective of 'double failure'", *Zhejiang Wanli University*, Vol 2, No. 1, pp. 11-15, 2014.

[18] Sun Y., "Analysis and suggestion on the status quo of the use and management of college funds," *Contemporary Accounting*, Vol 2, No. 1, pp. 66-67, 2016.

[19] Yang Q.R., "Analysis on the construction of rural public cultural service system in minority areas," *Caizhi*, Vol 2, No. 1, pp. 267-268, 2017.

A Study on the Influence Factors of User's Will to Use on Shared Residential Platform

Zhouyang Li[1], Weijia Li[2], Ziyang Liu[2*]
School of Management, Xinxiang University, Xinxiang, China
Department of Global Business, Kyonggi University,
Suwon, South Korea
[E-mail: lzy1990322@126.com, weijia1006@naver.com, morninglzy@hotmail.com]
*Corresponding author: Ziyang Liu

Abstract

In this paper, we study the factors affecting the user's willingness to use a shared residential platform. We study the effects of perceived usefulness, perceived ease of use, perceived entertainment, perceived risk on the willingness to use in a trusted intermediary.The data needed for empirical analysis is obtained by means of questionnaire investigation, and frequency analysis, element analysis, reliability analysis and structural equation analysis are carried out through SPSS V22.0 and AMOS V22.0 methods.

***Keywords*:** Sharing Economy. Perceived Usefulness. Perceived Ease of Use. Perceived Playfulness.

1. Introduction

Shared accommodation is a shared economy in tourism. It combines the idle tourist resources of the tourist site into one for the tourist consumers. It is the best choice for both supply and demand, and the supplier gains from the leasing of the idle resources [1]. In 2019, the share economy of taxi, restaurant and lodging accounted for 37.1 percent, 12.4 percent, 7.3 percent, and in the first five months of 2020, the share of the housing market in China dropped 72.1 percent year-on-year[2]. In order to provide scientific theoretical basis for the sustainable and healthy development of relevant enterprises in the shared housing environment, the sharing economy consumption pattern is moving forward in the direction of standardization and order [3].

2. Theoretical Background

Perception usefulness and ease of use comes from the Technical Acceptance Model (TAM), which is used primarily to understand the user's attitude to the new technology and when to adopt it.Davis (1985) defines the usefulness of perception as the extent to which people think the use of information systems increases their performance[4]. As the economy has developed in recent years, some scholars have used this variable to study users' acceptance of IT platforms. Based on this, the following assumptions are proposed:

H1: Consumer perception usefulness for trusta positive effect

H2 : Consumer perception ease of use has trusta positive effect

H3: Consumer perception of usefulness through trusthave a positive effect on one's willingness to use

H4: Consumer perception ease of use through trusthave a positive effect on one's willingness to

use.

H5: Consumer perception of the positive effect of entertainment on trust

H6: Consumer perception of entertainment through trust has a positive effect on willingness to use .
H7: Consumer Awareness Riskhave a negative effect on trust
H8: Perceptional Risk Through LettersIm has a negative effect on his willingness to use
H9: Consumer PerceptionTrust has a positive effect on willingness to use

Based on the above analysis, this paper constructs the research model shown in **Fig. 1**.

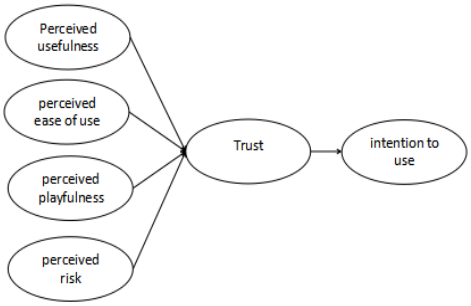


Fig. 1. research model

3. Research Method

Data were collected using survey questionnaire The subjects of the survey were those who had used a private housing reservation platform such as Airbnb.

4. Analysis and Results

4.1 Reliability Analysis

Reliability refers to the degree of consistency or stability of measured results. According to Nunnally's psychometric theory, Cronbach's α should be greater than 0.6, preferably greater than 0.7. The reliability of each variable and "item-deleted reliability" were analyzed using SPSS 22.0. The results are shown in Tab. 3. The scale reliability of each variable has met basic reliability requirement, but the reliability value was not satisfactory. The "item-deleted reliability" referred to the reliability that can be achieved after a test item was deleted. If the "item-deleted reliability" was greater than the original scale reliability and still there were no fewer than 2 test items in this structure variable after deletion, then this test item can be deleted. Therefore, SI5, SI6 and FC3 were deleted.

5. Conclusions and Limitations

Sensory usefulness and perceived ease of use significantly positive impact on trust and through trust positive influence on willingness to use.The user can meet the user's needs and demonstrate the platform's ability and ability to meet his or her promise, thus enhancing the user's confidence. When the user has confidence in the platform, he or she will choose the shared accommodation platform better and improve the user's willingness to use it by influencing the user's trust. Study the limitations of the sample.

References

[1] Li, Y., Ding, R., Cui, L., Lei, Z., and Mou, J., "The impact of sharing economy practices on sustainability performance in the Chinese construction industry," *Resources, Conservation and Recycling.* Vol 150, pp. 13-21, 2019.
[2] Ma, Y., Rong, K., Luo, Y., Wang, Y., Mangalagiu, D., and Thornton, T. F., "Value Co-creation for sustainable consumption and production in the sharing economy in China," *Journal of cleaner production.* Vol 208, pp. 1148-1158, 2019.
[3] Curtis, S. K., and Lehner, M., "Defining the sharing economy for sustainability. *Sustainability.* Vol 11, No. 3, pp. 567, 2019.
[4] Gupta, M., Esmaeilzadeh, P., Uz, I., and Tennant, V. M, "The effects of national cultural values on individuals' intention to participate in peer-to-peer sharing economy," *Journal of Business Research.* Vol 97, pp. 20-29, 2019.

The Impact of Creating Shared Value (CSV) Activities on the Image, Attitude and Purchase Intention in E-commerce

Huijie REN[1], Hao DAI[2*]
[1]School of PE, Taiyuan University of Technology,
Taiyuan, Shanxi, China
[2]Business Administration Graduate School, Kyonggi University
Suwon, South Korea
[e-mail: renhuijie1990@126.com, maple1314521@naver.com]
*Corresponding author: Hao DAI

Abstract

The purpose of this study is to examine the effect of shared value creation (CSV) activities of Chinese E-commmerce enterprises on the image of Chinese consumers on the E-commerce platform, the attitude on the E-commerce platform, the influence of image and consumer attitude on the intention of purchase and recommendation. And the image and attitude looked at the medium effect between CSV and the intention of purchase and recommendation.

The results of the study, first, the CSV activity of China's E-commersce has a significant effect of positive(+) Second, the CSV activity of China's E-commersce has a significant impact on attitudes. Third, the image of E-commerce has a significant positive effect on the intention of purchase. Fourth, the image of E-commerce has a non-significant effect on the intention of recommendation. Fifth, attitudes to E-commmerce have a significant positive effect on the intent of purchase. Sixth, attitudes to E-commerses have a non-significant effect on the recommendation intention. Seventh, the CSV activities of E-commers have a significant (+) effect on the recommendation intention.

Through this, we would like to recognize the importance of creating shared value of China's E-commerce company and provide important data in setting the CSV direction of China's E-commerce enterprise. Or to grasp the psychology of Chinese E-commerce enterprise consumers and provide strategies for the long-term operation of Chinese E-commerce enterprises.

***Keywords*:** CSV Activities, Image, Attitude, Purchase Intention, Recommendation Intention

1. Introduction

Enterprises should solve social requirements and problems such as environment and welfare, so they should make long-term investment to create sustainable income [1]. CSV is an activity that an organization can create benefits and solve social problems. That is, it can create economic value and social value at the same time, so as to improve the corporate social image. Therefore, CSV plays an important role in operation and attracts much attention. Recently, not only enterprises, but also infrastructure departments have taken advantage of CSV activities to improve their image [2][3]. Chinese E-commerce can also enhance users' visiting intentions through CSV.

2. Research Method

Based on the previous research above, the research model is shown in **Fig. 1**.

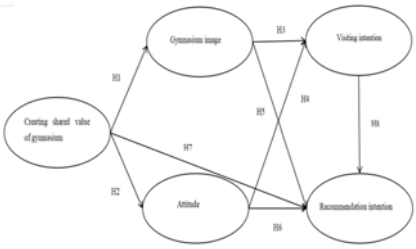


Fig. 1. Research Model

Research object

The purpose of this study is to verify the substance of this study by means of questionnaire survey of visitors who have visited E-commerce, China. A total of 400 questionnaires were distributed to the respondents, and 384 questionnaires were recovered in which 372 questionnaires were used in the final analysis except invalid answers.

3. Tourist satisfaction analysis

The results are as follows: firstly, the CSV activities of Chinese E-commerce have a positive (+) impact on the image of E-commerce. Secondly, CSV activities in Chinese E-commerce have a positive (+) impact on attitudes. Thirdly, the image of the gymnasium in China will have a positive (+) impact on the visiting intention. Fourthly, attitude will have a positive (+) influence on the visiting intention of E-commerce in China.

4. Conclusions and Limitations

According to the results, in order to enhance the good image of the gymnasium, it is more important to carry out more sustainable CSV activities than charity CSR activities while expanding the market field. In this era of fierce competition, the cognition of visitors to CSV activities of E-commerce can be used as an important means to enhance the image, attitude and visiting intention of E-commerce. Finally, the gymnasium should not only pursue profits, but also make the social and economic conditions and operation mode of the gymnasium continuously move towards a better direction through social contribution activities to the regional society in the rapidly changing external environment.

Through these activities, we realize the importance of creating shared value of Chinese E-commerce, and want to provide important information for setting CSV direction of Chinese E-commerce. In addition, I also hope to understand the psychology of visitors to Chinese E-commerce and provide strategic guidelines for the long-term operation of Chinese E-commerce.

References

[1] Barone, M. J., Miyazaki, A. D., & Taylor, K. A. (2000). The influence of cause-related marketing on consumer choice: does one good turn deserve another?. Journal of the academy of marketing Science, 28(2), 248-262.

[2] Crane, A., Palazzo, G., Spence, L. J., & Matten, D. (2014). Contesting the value of “creating shared value”. California management review, 56(2), 130-153.

[3] Everard, A., & Galletta, D. F. (2006). Effect of presentation flaws on perceived quality of on-line stores’ web sites what makes all the difference: The actual presence of a flaw or the perception of a flaw?. J. Management Inform. Systems, 22(3), 56-95.

KSII The 12th International Conference on Internet (ICONI) 2020.

The Influence of Service Quality of High-star Hotels and Online Word-of-mouth Marketing on Customer Loyalty

Yixuan Zhu and Kang Peng

Hunan Institute of Technology, Hu Nan, China
[e-mail: 85941138@qq.com, 13154400@qq.com]
*Corresponding author: Kang Peng

Abstract

The purpose of this research is to make suggestions for establishing customer loyalty and improving enterprise competitiveness of high-star hotels through analyzing the influence of service quality and online word-of-mouth marketing of high-star hotels. In this research, high-star hotel customers were used as the target and data was collected by questionnaires. After the questionnaire data was recovered, SPSS statistical software was used to analyze and process the statistical data, and AMOS software was used to analyze the equation model. The results shows that service quality of high-star hotels has a positive effect on customer perceived value; Service quality has a positive effect on customer loyalty; Online word-of-mouth marketing has a positive effect on customer perceived value; Customer perceived value has a positive effect on customer loyalty.

***Keywords*:** High-star hotels, Service Quality, Online Word-of-mouth marketing, Perceived value, Customer loyalty

1. Research background

Under the background of economic globalization, the tourism market of China continues to maintain a strong upward trend. The vigorous development of tourism has greatly stimulated the growth of demand for the hotel industry, which has brought both opportunities and challenges to the hotel industry. In recent years, the development of China's hotel industry has entered the stage of integration and upgrading, Among them, the supply of mid-range hotels has grown rapidly, ranking first in all grades of hotels, and high-star hotels have experienced a certain decline compared to previous years. On the one hand, as various capitals and enterprises have fought in the mid-range hotel market, promoting the transformation of the hotel market structure in China, and competition among hotel companies has become increasingly fierce; On the other hand, with the improvement of economic and cultural level， consumers tend to more mature and rational to consumption, and the diversification of the market also provides consumers with more choices.

2. Theoretical Background

Service quality refers to the sum of features and characteristics of a service that can meet specified and potential requirements. Lehtinen[1]divides service quality into entity quality, interaction quality and enterprise quality. Entity quality includes tangible facilities,

equipment, environment, products, etc.; Interactive quality refers to the interactive relationship between staff and customers; Enterprise quality refers to factors such as the reputation and overall image of an enterprise. Finnish scholar Gronnos[2] [3].

Therefore, this research will quantitatively measure service quality from six dimensions: reliability, responsiveness, assurance, tangibility, empathy, and service remediation capability.

3. Research methods

Variables in the model were measured using a multi-item measurement, and the items are measured by Liker7 component table. Among them, "1" means "Completely dissatisfied", "2" means "Mostly dissatisfied", "3" means "Somewhat dissatisfied" , "4" means "Neither satisfied or dissatisfied", "5" means "Somewhat satisfied","6" means "Mostly satisfied", and "7" means " Completely satisfied".

During the pre-investigation phase, 30 questionnaires were distributed, and the contents of the questionnaires were further improved through the information feedback of 30 customers. A total of 240 questionnaires were distributed during the formal investigation phase and 226 questionnaires were recovered, among which 16 invalid questionnaires were excluded, and 210 valid questionnaires were finally recovered with the actual recovery rate being 87.5%. According to the results of valid questionnaires, SPSS 22.0 was used for reliability analysis, factor analysis and structural equation model analysis.

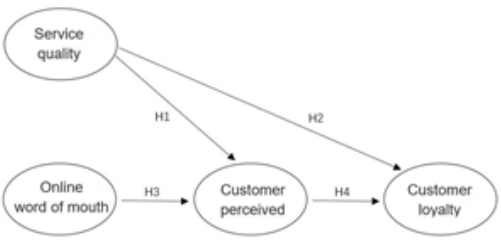


Fig. 1. Research Model.

4. Conclusions

Based on SERVPERF model and referring to previous theoretical research result, this paper adds a remedy dimension that affects the service quality of high-star hotels. Through empirical analysis and hypothesis verification, besides traditional factors such as tangibility, reliability, responsiveness, assurance and empathy, the remedy also impose significant impact on the perceived service quality of high-star hotel customers. Therefore, all employees in the high-star hotels, from the hotel's senior management to the front-line service staff, should shape a correct service remediation advice. A service remediation system from remediation early warning, on-site remediation to remediation feedback should be established.

The service quality of high-star hotels imposes positive impact on the customer loyalty; and the service quality could impose indirect impact on the customer loyalty through the intermediary effect of customer perceived value. To the development of high-star hotels, maintaining the customer loyalty should start from improving service quality and customer perceived value.

Under the era background of "Internet +", Online word-of-mouth marketing has been a popular marketing method and its value is becoming increasingly apparent. Online word-of-mouth marketing makes influence on customer loyalty through the customer perceived value. Therefore, high-star hotels should focus on the strategy of Online word-of-mouth marketing.

References

[1] J. B. Baty, and R. M. Lee, "Intershop: Enhancing the vendor/customer dialectic in electronic shopping," Management Information Systems, Vol 11, No. 4, pp. 9–31, 1995.

[2] D. M. Szymanski and R. T. Hise, "E-satisfaction and Initial Examination," Journal of Retailing, Vol 76, No. 3, pp:309-322. 2000.

[3] Z. Jinxiang, J. Fanghui, and W. Lisheng, "Customer Satisfaction in E-commerce: an Exploration of its Antecedents and Consequences," IEEE International Conference on Management of Innovation & Technology IEEE, 2006.

An Analysis of Chinese Local College Students Employment Choices and Factors Influencing the Choice

Shuo Liu
Xinyang Normal University, China
[e-mail: 335481794@qq.com]
*Corresponding author: Shuo Liu

Abstract

The purpose of the paper is to make a comprehensive evaluation of the employment of local college students, to explore the selection of employment regions for college students, and to help regional college students establish a correct employment concept to accelerate the recruitment and employment and teaching management to promote urban development decision-making. The analysis of this paper is concluded by questionnaires and in-depth interviews. According to the 69000 college students in Xinyang, 1013 questionnaires were distributed and 1013 questionnaires were recovered. 800 questionnaires were effective. This paper ranks the importance order for the employment influencing factors and forecasts the choice of employment place by using descriptive statistical method text analysis and random forest method. It is concluded that the college students' expected salary is the most important among 26 factors surveyed, followed by employment place and its household registration policies. Low cost of living and parents' education level ranked third and fourth respectively.

***Keywords*:** R in action; text analysis; choice of employment place; random forest method

1. Introduction

Making choice of employment area is an important part of employment behavior. It is of great practical significance and theoretical value to explore the choice of employment region and its influencing factors to guide college students to choose their jobs, get employed reasonably and help colleges improve employment services and make right decisions for city's development.

It's not that easy for college students to find a job currently .On the one hand, the number of colleges graduates is increasing year after year, 5.7544 million in 2010, and 7.533 million in 2018(**Fig. 1**).This situation leads to the supply of college students' employment market exceeds the demand, making college students hard to find a job. On the other hand, what students learn in college is not what society needs. Moreover, college student's career plan is not sufficiently organized and their career concept is out-of-date, which make finding jobs a hard thing. Difficulty in seeking jobs mainly reflected in the difficulty of selection of students' employment region.

2. Design and data analysis

The sample size is 843 due to the matter of sampling the quota in the non -probabilistic sampling used in this analysis. The proportion of male students in effective questionnaire was 25.13%, that of female students was 74.87% .Male and female ratio of the samples was basically the same as that of the overall male and female number .The distribution of the grade of sample is about 1:1:1:1,which is basically the same as the overall distribution. Therefore, the sample distribution is more consistent with

overall distribution and samples are representative.

3. Empirical Analysis of the willingness of college students to choose employment

According to the study results, 86％ of the students are non-only children in Xinyang's universities and the only-children students accounts for 14％. The families with middle level economical condition take the largest part,occupy 53.8％. The families with lower economical condition and the wealthy families account for 24.7％ and 21.3％ respectively. The surveyed parents generally do not receive a high education and most of them have a schooling of high school education or do not get high school education. The parents with bachelor degree or postgraduate diploma only take up 5％. Further more the research shows that 0.5％students lack communication with their parents. However, most of them live in a harmonious family (table3).

4. Conclusions

Consequently, the college should vigorously develop the employment services and improve their guidance and service to college graduates in their search for jobs and career opportunities. According to social development and employment, colleges need to analyze the industry circumstance, cooperate with enterprises and institutions actively, take advantage of the existing resources and exploit uncovered resources, and broaden employment channels for graduates. As parents, they should communicate with their more often, do not expect too much of children, help them analyze the current employment situation and provide appropriate guidance of employment. All of these aim to contribute to promote employment space and career prospects of college students, speed up the formation of characteristic industries in accordance with local condition. Collages should form the competitive advantage of differentiated industries, and make the colleges more attractive to professionals in specific fields.

References

[1] M. E. Menon, N. Pashourtidou and A. Polycarpou, et al. “Students’ expectations about earnings andemployment and the experience of recent university graduates: Evidence from Cyprus,” International Journal of Educational Development, Vol 32, No.6, pp. 805-813, 2012.

[2] V. Vasile. “Crisis impact on employment and mobility model of the Romanian university graduates," Procedia Economics and Finance, Vol 3, pp.315-324, 2012.

[3] P Rérat. “The selective migration of young graduates: Which of them return to their rural home region and which do not?" Journal of Rural Studies, Vol 35, pp.123-132, 2014.

[4] Chen Chunchao and Qi Wanning. “An investigation on the influencing factors of College students' Employment region Choice: a case study of Nanjing University students," World of Labor Security, Vol 18, pp. 14-15, 2018.

[5] Song Guokai, Wu Tian and Li Dong. “The influence of Family status, Human Capital and Policy on College students' Employment willingness at the Grass-roots level," Journal of Xi'an Jiaotong University (Social Science Edition), Vol 37, No.5, pp. 71-78, 2017.

[6] Wu Junhui. “An Analysis of Employment intention of Agricultural College students in Xinjiang," Xinjiang Agricultural University, 2016.

[7] Xu Yingmei and Xia Lun. “Analysis of influencing factors of subjective well-being of Chinese residents: a Comprehensive Analysis Framework," Journal of Central South University of Finance, Economics and Law, Vol 2, pp.12-19, 2014.

Deep Neural Network-Based Deep Packet Inspection for Improving Traffic Steering in Real-Time IoT Network

Sa Math[1], Prohim Tam[1] and Seokhoon Kim[2]*
[1]Department of Software Convergence, Soonchunhyang University, 22, Soonchunhyang-ro, Sinchang-myeon, Asan, 31538, Rep. of Korea
[2]Department of Computer Software Engineering, Soonchunhyang University, 22, Soonchunhyang-ro, Sinchang-myeon, Asan, 31538, Rep. of Korea
[e-mail: mathsa98@gmail.com, prohimtam@gmail.com, seokhoon@sch.ac.kr]
*Corresponding author: Seokhoon Kim

Abstract

With the rapid growth of intelligent devices and communication technologies, the 5G network environment has become more heterogeneous and complex in terms of management and orchestration. 5G architecture requires supportive technologies to handle the existing challenges for improving the Quality of Service and the Quality of Experience performances. Among many challenges, traffic steering is one of the key elements which requires to critically develop an optimal solution for smart guidance, control, and reliable system. Mobile Edge Computing, Software-Defined Networking, Network Slicing, and Machine Learning play essential roles to complementary develop a flexible and extensible flow rules management in this potential aspect. The proposed system provides an accurate flow recommendation, a centralized control, and a reliable distributed connectivity based on packet condition inspection. By this implementation, the packet is classified separately and recommended to request from the optimal destination with matched preferences or conditions. Intelligent steering for network communication traffic leads to a variety of beneficial factors for improving real-time Internet of Things communication performance.

***Keywords*:** Deep Neural Network, Traffic Steering, Deep Packet Inspection, Mobile Edge Computing, Software-Defined Networking, Network Slicing

1. Introduction

With the current development and implementation of newly modernized communication technologies (e.g. Intelligent Internet of Things devices), network transmission traffic in the 5G network is increasing notably. The growth provides many challenges in terms of management, reliability, complexity, optimization, and security in the heterogeneous network environment which associates with various types of devices, protocols, data centers destination, application (e.g. conversation, gaming, streaming, interactive, background), and processing mechanisms. By this diversification, the network architecture has to be flexible and extensible for handling the potential manifold issues [1]. In these recent years, many researchers and

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2020-2015-0-00403) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation), and this research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2020R1I1A3066543).

organization have struggled to develop supportive advanced enabler technologies to tackle the existing challenges such as Mobile Edge Computing MEC), Software-Defined Networking (SDN), Network Slicing, Network Function Virtualization (NFV), massive MIMO (multiple-input and multiple-output), the emergence of 5G mmWave (millimeter wave), D2D (device-to-device) communication, etc. [2].

Related Works

MEC is a state-of-the-art paradigm of distributed communication and computation server availability at the edge of the network which extracts the significant resource allocation from the central cloud server to the edge server including computing capacity, processing memory, communication bandwidth, and other energy resource reservation. From centralized computing, network control, and storage capacity, Mobile Cloud Computing (MCC) drives complex and tremendous growth of big data communication [3-4]. With MEC supports, the concerned action moves toward network edges and provides faster speed, more available storage space, greater computing power capability, and higher possibility of developing or implementing new Internet of Things (IoT) technologies. MEC has been merged with ML algorithms to contribute minimal latency in offloading decisions, allocating the resource, controlling overhead management, and handling corporate-server communication services [5-6]. Without the implementation of appropriate algorithms, MEC faces several issues including complicated configured administration, big data processing, and complex high-level unlabeled parameters. Under the same circumstance, MEC traffic steering also requires optimal ML algorithms to detect and classify diversified traffic packet statuses for an accurate recommendation and further navigation.

3. Deep Neural Network-Based Deep Packet Inspection Approach

In 5G communication network environment, various packets in terms of destination application, communication protocols, packets types, different Quality of Services parameters, and other packet characteristics are transmitted and received. With multiple incoming packets which have different conditions at the network devices, the complexity and insufficient steering are happened.

The packet classification techniques are commonly used by port-based approach, DPI, and ML-Based. For port-based technique, the uses of UDP and TCP port numbers determined the main application of the traffic, but presently, the dynamic ports are used by many other applications or platforms which is the main cause of ineffective port-based capability. Therefore, DPI is better and more reliable method to classify and detect the traffic type.

Nevertheless, there are remaining challenges such as long time-consuming process, pattern-constrained, and high energy-consuming requirement. Thus, ML and DPI convergence becomes a considerable approach to handle these issues. With sufficient algorithms, the proposed approach manages to gather big data of packets' condition and flow pattern as training dataset to develop the model. ML-Based DPI requires the understanding of packet details with pre-determined pattern analysis (input datasets) for identifying the destination that the flow rules forward to. Because of the rapid growth of application services, the pattern of packet becomes more challenging and problematic for inspection purposes. Big collection of packet traffic flows is required to complete for processing an accurate and reliable ML-Based DPI. Figure 1 illustrates the brief processing flow of selecting Deep Neural Network (DNN) algorithm to apply for packet detection for outputting a reliable clustering result for intelligent traffic steering configuration in OpenFlow SDN controller.

DNN uses multiple Neural Networks with multiple hidden layers between input and output layers. By using Multi-Layer Perceptron (MLP) algorithm, each neural part of hidden layers plays an important role to contain the value and analyze the data of the previous layer's neural condition for given weighted metric summary report. MLP allows the whole algorithm to enable non-linear or unlabeled packet flow patterns in real-time communication. After the model training stage, DNN is capable of generating an accurate and reliable result of packet classification, server

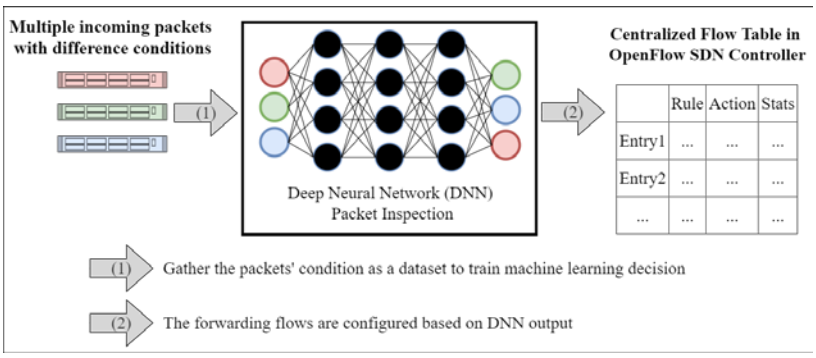


Fig. 1. The proposed Deep Neural Network-Based Deep Packet Inpection Approach

recommendation, or application detection purposes. Thus, the DNN result is crucial and cooperative for the next stage of SDN controller configuration. As discussion above, SDN control plane consists of centralized programmable controller (e.g. POX, NOX, Ryu, Floodlight) which allows the proposed system to improve edge traffic steering by configuring and setting the flow table critically by using OpenFlow protocol to interact with the data plane. By detecting the rule status, network function supports virtually for matching up the connectivity which has similar preference to the historical dataset. At the final stage, edge traffic steering keeps synchronizing the forwarding rules with SDN controller continuously and updating the fault-tolerance or high availability discovery. Concurrently, DNN stores the historical flows and statuses for better understanding of the surrounding network environment and advances the next action of the next state for better calculation and reliability.

5. Conclusions

The effective traffic handling methods are required for the overcome the massive IoT traffic to meet the perspective of 5G communication technology. This paper presents the effective real-time IoT services handling based on deep neural networks for improving end-to-end network quality of service. The proposed method is focused on the classification of incoming traffic into different groups and different capacity of serving servers. The scheme has two big contributions, balancing the incoming traffic to meet the capacity of MEC servers, and inspect the different QoS requirements of IoT packets and handling based on the recommended machine learning and deep learning algorithms. The future research extending, we will experiment with the second part of the proposed scheme.

References

[1] P. Wang, C. Yao, Z. Zheng, G. Sun and L. Song, "Joint Task Assignment, Transmission, and Computing Resource Allocation in Multilayer Mobile Edge Computing Systems," *IEEE Internet of Things Journal*, Vol. 6, No. 2, pp. 2872-2884, 2019.

[2] S. Kekki et al., "MEC in 5G networks," *ETSI white paper*, No. 28, pp. 1–28, 2018.

[3] V. Nguyen, A. Brunstrom, K. Grinnemo and J. Taheri, "SDN/NFV-Based Mobile Packet Core Network Architectures: A Survey," *IEEE Communications Surveys & Tutorials*, Vol. 19, No. 3, pp. 1567-1602, 2017.

[4] X. Wang, Y. Han, V. C. M. Leung, D. Niyato, X. Yan and X. Chen, "Convergence of Edge Computing and Deep Learning: A Comprehensive Survey," *IEEE Communications Surveys & Tutorials*, Vol. 22, No. 2, pp. 869-904, 2020.

[5] T. K. Rodrigues, K. Suto, H. Nishiyama, J. Liu and N. Kato, "Machine Learning Meets Computation and Communication Control in Evolving Edge and Cloud: Challenges and Future Perspective," *IEEE Communications Surveys & Tutorials*, Vol. 22, No. 1, pp. 38-67, 2020.

[6] J. Xie *et al.*, "A Survey of Machine Learning Techniques Applied to Software Defined Networking (SDN): Research Issues and Challenges," *IEEE*

KSII The 12th International Conference on Internet (ICONI) 2020.

Traceability Issues in Segregated Learning with Decentralized Incentives

Sandi Rahmadika[1] and Kyung-Hyune Rhee[2]
[1]Interdisciplinary Program of Information Security, Graduate School PKNU
[2]Department of IT Convergence and Applications Engineering
Pukyong National University, Busan, South Korea
[e-mail: sandika@pukyong.ac.kr, khrhee@pknu.ac.kr]
[2]Corresponding author

Abstract

This paper presents several concerns in adopting blockchain-based incentives in a segregated learning system. The main objective of segregated learning is to preserve privacy for the clients after training. It enables numerous devices to conduct training in a global model by using clients' data privately. The contributed clients are incentivized by relying on the blockchain smart contract features. However, the *two-phase commit* designed between parties breaks the privacy. By nature, the devices frequently send their updated gradient values back to the server that makes the information values leak to the malicious parties. Moreover, malicious parties with certain assumptions can trace the address of the data. Therefore, in this paper, we observe essential points related to the traceability concerns in segregated learning with blockchain-based incentive.

***Keywords*:** Blockchain, privacy, segregated learning, smart contract, traceability issues

1. Introduction

Presently, the paradigm of using a concentered network topology approach as the core backbone in controlling the machine instruction commands has been positively refashioning since the emergence of Bitcoin, Ethereum blockchain and federated learning [1]. By design, blockchain is irreversible data and tamper-proof chain of transaction records that can solve the dispute between parties in the same blockchain network.

Segregated learning and training in the artificial intelligence (AI) area has been extensively researched recently. Segregated AI provides a breakthrough in the deep learning area since it converts the concentered form of raw data into a separated model [2]. The raw data possessed by clients are never leaving the devices. Thus, the severe issues of privacy in concentered learning can be tackled by design.

Ethereum smart contracts are self-executing contracts with immutable data records that can be a plausible solution in propagating incentives within segregated learning schemes. However, the *two-phase commit* designed between users and model providers through the blockchain environment could break the users' privacy that is contrary to the principal objective of segregated learning. Hence, by adopting the inference attacks, the observer can infer and link the presence of specific data features dataset.

In this paper, we outline several fundamental points related to the concept, design and traceability issues of implemented schemes. Segregated learning is assumed to be implemented for highly sensitive data such as health-related data since the gradient values from users are exposed publicly through the smart contract. With certain premises, the observer can

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. NRF-2018R1D1A1B07048944) and partially was supported by the Republic of Korea's MSIT (Ministry of Science and ICT), under the High-Potential Individuals Global Training Program) (2020-0-01596) supervised by the IITP (Institute of Information and Communications Technology Planning & Evaluation).

adopt the active and passive inference attack to impose training data through gradients value. Eventually, it breaks privacy. The transactions are also traceable. Therefore, we put forward these points to be investigated in this paper.

2. Merging Blockchain Technology and Segregated Learning

2.1 Collaborative Intelligence

One of the most prominent benefits of utilizing blockchain technology is not involving a middleman nor intermediaries in managing the transactions [3]. This feature is also useful for segregated intelligence where data is no longer concentrated. The data is scattered among the user devices in the same application. In the segregated learning algorithm, there is a set of users with a distinct local dataset.

$$f(w) = \frac{1}{n}\sum_{i=1}^{n} f_i\,(w) \tag{1}$$

$$\sum_{n=1}^{n} \frac{c_n}{c} w_{t+1}^{n} \tag{2}$$

Intuitively, $fi(w) = (xi, yi, w)$ can be defined as a loss of the prediction on example (xi, yi) which is managed by model parameters w as shown in (1). It also can be interpreted as IoT devices send gradients or parameters $\Delta w1+\Delta w2+\Delta w3+\cdots+\Delta wn$ to the cloud server, which is partitioned homogenously [4]. Finally, the aggregated server computes the model received and applies it to the new parameters as defined in (2).

2.2 Segregated Learning with Ethereum Smart Contract

A blockchain-based loyalty rewards system can reduce management costs with smart contracts that are secure, trackable, transparent, and lower costs. Ethereum smart contracts can be leveraged further, such as merging with AI, as shown in **Fig. 1.** This combination can provide a more resilient and efficient path for an unconcentrated system.

Fig. 1 shows a smart contract that manages the training of AI agents with an interpretable output. As the transition model towards the blockchain space, this could become a critical tool in offering more reliable and secure deconcentrated applications (DApps).

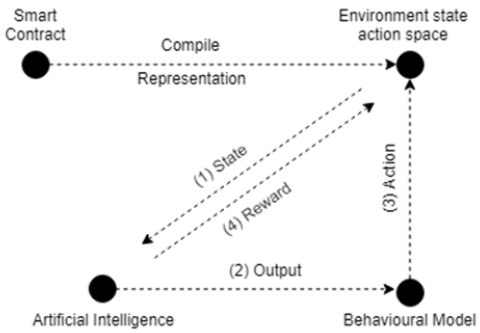


Fig. 1. High-level structure of the AI smart contract

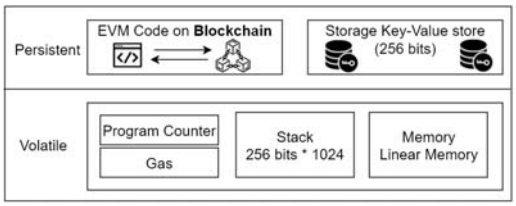


Fig. 2. Ethereum Virtual Machine (EVM)

Ethereum platform preserves a decentralized ecosystem for developers to create products using the Ethereum Virtual Machine (EVM) which is powerful and embedded within each full blockchain network as shown in **Fig. 2.** The smart contracts bytecodes are executed through Ethereum Virtual Machine. Interacting with the EVM via smart contracts is likely more costly than traditional servers. However, numerous use cases are favored using EVM rather than conventional servers.

3. Traceability Concerns and Outlook

Private segregated learning. The clients are data owners who hold a large amount of individual training data, while the AI provider preserves the deep learning model that can be accessed by authorized clients. The updated gradient are collected by the provider regularly, which later to be employed to calculate the final aggregation value. This process is repeated as long as necessary.

The AI provider also provides revenue through smart contracts (via EVM) for those who have been successfully proven contributing to improving the global models (using their private resources without exposing the actual data). A two-phase commit transaction is adopted during transactions.

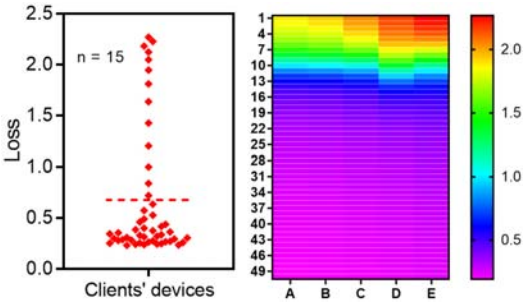


Fig. 3. Distribution points of average loss

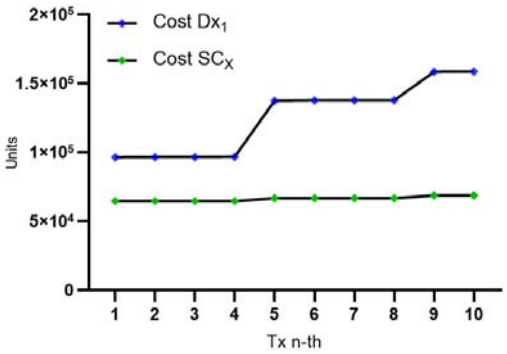


Fig. 4. Gas comparison (AI providers and clients)

The distribution points of an average loss of our scheme are shown in **Fig. 3**. We set several devices to conjointly building an AI model. The average loss for every device is almost identical, which is around 0.671 up to 0.786. Precisely, the number of devices involved in our scheme slightly affects the loss.

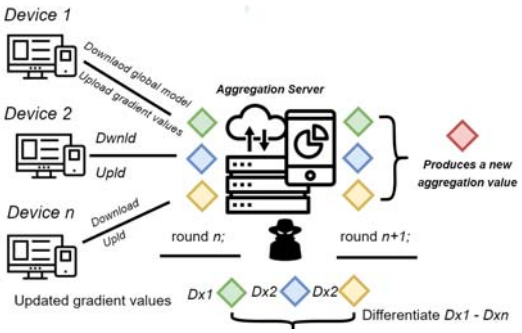


Fig. 5. The inference attacks

Fig. 4 shows the comparison amount of gas used between AI provider *SCx* and clients *Cx* on Ethereum smart contracts. The AI provider consumes gas stably for every transaction, while the clients consume more gas depend on their contribution (arbitrarily stated in the contracts). The clients' contribution affects the gas used.

Traceability concerns. When the clients deploy the transaction that consists of a cipher to encrypt the information to the AI provider, the observer can impose the dataset knowledge by adopting active and passive inference attacks as shown in **Fig. 4**. Yet, the performance of the adversary decreases with an increasing number of users as elaborated in [5]. In short, as batch sizes increase, this type of attack produces more false positives. More promptly, the size of the batch influences the precision of the adversary. The larger the size of a batch, the lower the precision from the adversary's side.

4. Conclusions

We have presented a segregated learning scheme combines with blockchain technology as an incentive mechanism. The objectives are to provide the essential points in adopting segregated learning with Ethereum smart contract as a platform for distributing incentives. The performance of segregated learning satisfies the design goal in general. However, the output of transactions can be traced publicly. Therefore, additional protocols are a necessity to be adopted. For the journey of the future, we extend the blockchain merits to replace the role of the aggregation server. Finally, the system becomes fully decentralized without relying on the central server to calculate the gradient values.

References

[1] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated Learning: Strategies for Improving Communication Efficiency," pp. 1–10, 2016.

[2] G. B. Giannakis, Q. Ling, G. Mateos, I. D. Schizas, and H. Zhu, "Decentralized Learning for Wireless Communications and Networking," 2016.

[3] G. A. Montes and B. Goertzel, "Distributed, decentralized, and democratized artificial intelligence," *Technological Forecasting and Social Change*. 2019.

[4]W. Wu, L. He, W. Lin, R. Mao, C. Maple, and S. A. Jarvis, "SAFA: a Semi-Asynchronous Protocol for Fast Federated Learning with Low Overhead," *IEEE Trans. Comput.*, 2020.

[5] L. Melis, C. Song, E. De Cristofaro, and V. Shmatikov, "Exploiting unintended feature leakage in collaborative learning," in *Proceedings - IEEE Symposium on Security and Privacy*, 2019.

KSII The 12th International Conference on Internet (ICONI) 2020.

Optimizing Actuators Deployment with Multi-layer particle swarm optimization for wireless sensor and actuator networks

Yamin Han and Heejung Byun
Department of Information and Telecommunications Engineering, The University of Suwon — South Korea
[e-mail: yaminhan11@gmail.com, heejungbyun@suwon.ac.kr]
*Corresponding author: Heejung Byun

Abstract

In wireless sensor and actuator networks (WSANs), the network lifetime is an important criterion to measure the performance of the WSAN system. The network lifetime is mainly affected by the residual energy of sensor nodes. The energy of sensor nodes is limited and it is crucial to make energy consumption efficient. Moreover, in large-scale networks with a large number of sensor nodes, multiple sink nodes should be deployed, not only to increase the manageability of the network, but also to reduce the energy dissipation at each sensor node. Actuators introduced by WSAN can be regard as multiple sinks and deployed in the network system to collect data from their respective surrounding sensor nodes. An improved multi-layer particle swarm optimization algorithm is adopted to optimize the deployment of multiple sinks to maximize the coverage of sink to sensor nodes and reduce the energy consumption of sensor nodes. Simulation results show that this method is effective for improving the coverage of sink to sensor nodes and reducing the energy consumption of sensor nodes.

***Keywords*:** Wireless sensor and actuator networks, coverage, energy consumption, particle swarm optimization

1. Introduction

Wireless sensor and actuator networks [1] consist of sensor nodes and actuator nodes. Sensor nodes with limited energy and the short transmission range; actuators with rich energy and longer transmission capabilities. Sensor nodes can detect events from the monitored region, then they process and transmit the event data to their nearby actuators. However, the number of sensor nodes is relatively large and the number of actuators is limited in a WSAN system. Also, the energy consumption of sensor nodes is directly related with the distance between the sensor node and its nearest actuator. Therefore, it is necessary to research how to deploy a limited number of actuators according to the distribution of sensor nodes.

To solve the problem of actuators deployment, this research adopts an improved multi-layer particle swarm optimization algorithm to optimize the deployment of actuator nodes, so as to maximize the coverage of actuators to sensor nodes and reduce the energy consumption of sensor nodes.

2. Problem description and system modeling

We assume N number of sensor nodes are randomly scattered in the monitoring region R.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2020R1A2C1004390).

All the sensor nodes are homogeneous, i.e., all the sensor nodes have same communication distance and sensor distance. Actuators can discover each other and collaborate with other actuators in the monitoring region through wireless communication, and they can also gathering data from sensor nodes. Sensor nodes send data and the actuators receive data consume energy, and the calculation model of the energy formula uses the calculation method [2].

For the coverage of actuators to sensor nodes, the detection probability model [3] is used to calculate whether a sensor node can be covered by actuators

3. Optimizing Actuator Deployment with Multi-layer particle swarm optimization

2.1 Standard Particles Swarm Optimization

The particles swarm optimization (PSO) algorithm is inspired by the behavior of birds to establish a dynamic model that can be employed in optimization. In 1998, Shi and Eberhart [4] introduced inertia weight into the original PSO algorithms, which is called standard PSO (SPSO). Each particle represents a potential solution to an optimization problem, and all particles form a population flying in a D-dimensional search space. The flying direction and distance of a particle is based on its velocity that is influenced by the best solution found by the population *Gbest* and the best solution *Pbest* found by itself. In each iteration step, updating the velocity v_i and position x_i of a particle in the *i*-th dimension are as follows:

$$v_i = wv_i + rc_1(Gbest - x_i) + rc_2(Pbest - x_i)$$

$$x_i = x_i + v_i$$

where *r* is a random value within [0,1]. c_1 and c_2 are acceleration constants which control the influence of the social and cognitive components.

PSO can effectively optimize the parameters. However, for complex and high dimensional problems, it is prone to fall into local optimal or premature convergence.

2.2 Multi-layer Particle Swarm Optimization

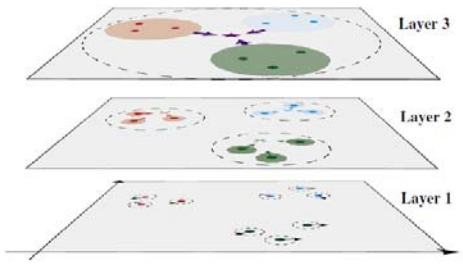


Fig. 1. Composition of particles

The proposed multi-layer particle swarm optimization (MLPSO) [5] can improve the performance of traditional PSO by increasing the two layers of swarms to multiple layers. The MLPSO strategy increases the diversity of searching swarms to improve its performance when solving complex problems. The sketch of layered view of a three layers MLPSO is shown in **Fig. 1**. The velocity update of a particle is determined by the accumulated information from all layers. The velocity formulas of the *i*-th dimension of a particle is shown as follow:

$$v_i = wv_i + \sum_{j=1}^{m} c_j r(Pbest_{i,j} - x_i), c_j = \frac{c}{m}$$

where *m* is the layer of the particle swarm, *c* is the acceleration constant.

2.3 Optimizing the Actuators Deployment with MLPSO

Assuming *K* number of actuators in WSAN, there will be 2n parameters where every two parameters correspond to the location of an actuator to be optimized. A particle can be encode by (x_1, y_1, x_2, y_2,..., x_n, y_n), where (x_i,y_i) is the potion of the *i*-th actuator, as shown in **Fig. 2**.

x_1	y_1	x_2	y_2	...	x_K	y_K

Fig. 2. Composition of particles

The goal of this work is to optimize the actuators deployment, taking into account both coverage and energy consumption.

4. Simulations and Results

To verify the performance of MLPSO method for actuators deployment in WSAN, the simulation is done using MATLAB 2018. MLPSO method is compared to the standard PSO (SPSO) [4]. The simulation settings are as below: 300 sensor nodes are distributed in a $400m \times 400m$ two-dimensional monitoring region; MLPSO and SPSO methods iterate 1000 times respectively; the number of collectors K is 20. Furthermore, to reduce statistical error, all methods are simulated 30 times under the same conditions, and the average performance is given as the last results.

Table I shows the coverage rates, and energy consumption rates obtained by MLPSO and PSO when optimizing actuators. The MLPSO gets the highest coverage rate and the lowest energy consumption rate. Since MLPSO searches for the optimal solution in multiple layers, which prevents the method from falling into the local optimum.

	Coverage Rate	Energy Consumption Rate
SPSO	81.13%	2.38%
MLPSO	96.50%	2.03%

4. Conclusions

In this paper, the problem of actuators deployment is studied. Actuators are deployed based on the distribution of all sensor nodes in the WSAN system, which considers the coverage rate of actuators to sensor nodes and the energy consumption rate of sensor nodes. The actuators deployment problem is formulated as an optimization problem optimized by an improved MLPSO method. The simulation results show that the MLPSO method can effectively optimize actuators deployment by increasing the coverage rate of actuators to sensor nodes and the energy consumption rate of sensor nodes.

References

[1] N. Primeau, R. Falcon, R. Abielmona, and E. M. Petriu. "A review of computational intelligence techniques in wireless sensor and actuator networks." *IEEE Communications Surveys Tutorials*, 20(4):2822−2854, Fourthquarter 2018.

[2] Jorio, Ali, et al. "An energy-efficient clustering routing algorithm based on geographic position and residual energy for wireless sensor network." *Journal of Computer Networks and Communications 2015* (2015).

[3] E. Onur, C. Ersoy, H. Delic, and L. Akarun. "Surveillance wireless sensor networks: Deployment quality analysis." *IEEE Network*, 21(6):48−53, November 2007.

[4] Y. Shi and R. Eberhart. "A modified particle swarm optimizer." In 1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98TH8360), pages 69−73, May 1998.

[5] Lin Wang, Bo Yang, and Yuehui Chen. Improving particle swarm optimization using multi-layer searching strategy. Information Sciences, 274:70 − 94, 2014.

Optimized Encoder Settings in Vehicle Network Environment

Byoungman An and Youngseop Kim
Electronic and Electrical Engineering, Dankook University
152, Jukjeon-ro, Suji-gu, Yongin-si, Gyeonggi-do, 16890, Korea
[e-mail: yesbman@naver.com, wangcho@dankook.ac.kr]
*Corresponding author: Yeongseop Kim

Abstract

This paper presents the latest Ethernet standardization of in-vehicle network and the future trends of automotive ethernet technology. The proposed system provides a design and optimization algorithm of in-vehicle networking technologies related Ethernet Audio Video Bridge (AVB) technology. We present the latency of AVB network as well as purpose the optimized design of the Ethernet network system for automotive. A proposal of Reduced Latency of Machin to Machine (RLMM) plays a significant role in reducing the delay between components. The RLMM approach on realistic test cases showed that there was a delay reduction about 41%. It is expected that the optimization method for the actual automotive environment can greatly shorten the time period in the design and development process. The results achieved from the experiments on the latency present in each function are reliable because average values are obtained through repeated actual tests for several months. It will greatly benefit the industry since analyzing the latency between each function in a short period of time is very meaningful.

***Keywords*:** Automotive Multimedia, Audio Video Bridge, RLMM, In-vehicle Network, AVB

1. Introduction

Automotive application has many differences in the use-cases and technical requirements as compared with various types of application; thus, the new audio / video transmission technologies as Ethernet Audio Video Bridge (eAVB) are emerged. [1] This presentation starts with the background for the technological introduction into automotive network based on the features of Ethernet and in vehicle requirements. The time sensitive video and audio data streams in an automotive bridged local area network (LAN) requires the selection of specific features that are specified in a number of various standard because the media is increasingly transported in automotive environment. Streaming audio/video with interactive applications over networks need to have comparable real time performance with legacy analog features. Vendors and users significantly interest in defining a common method for handling real time audio/video. The simple audio/video transport over multiple IEEE 802 network types will achieve operational benefits. IEEE Ethernet AVB (Audio Video Bridge) technology and standards provide transport mechanisms for time sensitive audio/video streams, it lessens the effort of producing inter working units among various networks. [2] A plenty of technology has been aimed to research the low latency of distributed real time network in vehicle. [3] Several study studies present various deployment and test configurations from simple networks with two end nodes and one switch to more complex networks with multiple endpoints and a multitude of switches between them. [4,5,6,7,8]

This research was supported by a research grant from Dankook University

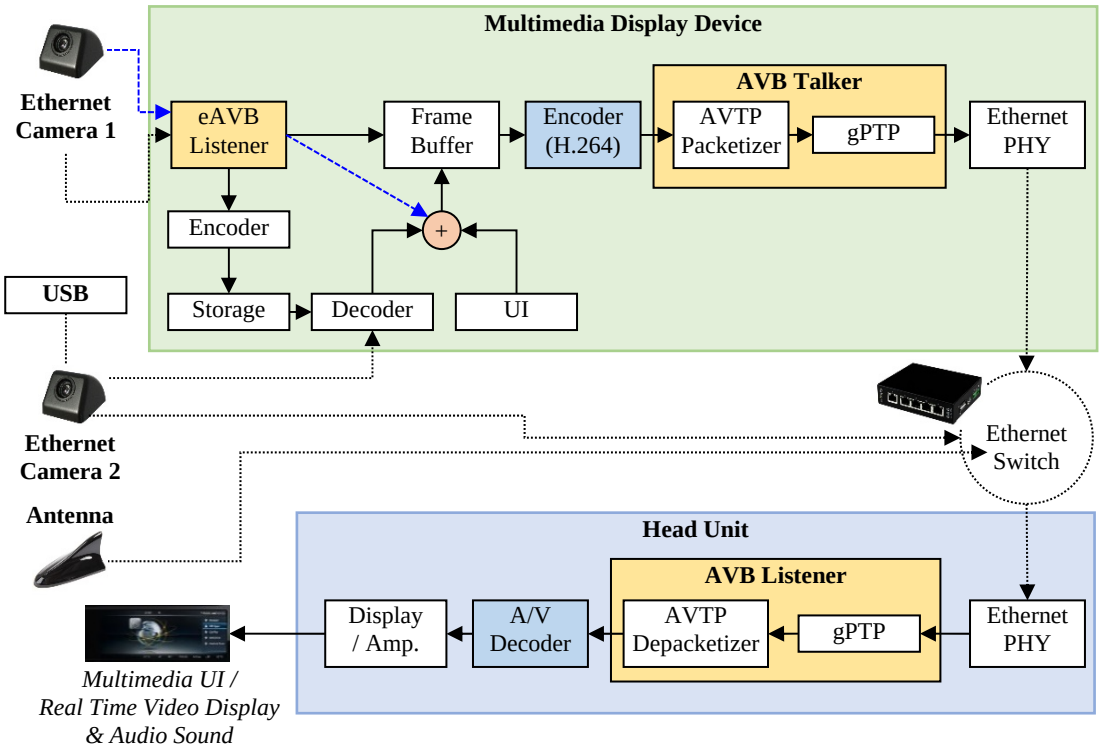


Fig. 1 An example of machine to machine latency in vehicle environment between multimedia display device and head unit, blue dotted line is a real-time camera video transmission without recording.

This paper aims to provide a new architecture of using an integrated framework with the eAVB networking. This paper is structured as follows: Section 2 describes the eAVB standards with the major concept. Section 3 presents our preliminary work as well as purposed idea. Section 4 presents the experimental results from purposed designs. Finally, Section 5 concludes the paper before proposing the future of study.

2. Ethernet Audio Video Bridge

Ethernet Audio Video Bridge (eAVB) is a fairly new technology with active work, but small number of commercial devices which are switches or particular endpoints. As endpoints, eAVB capable speakers, digital signal processors, and cameras are valid. To be appropriate for greatly integrated automotive system where additional data can often allow multiple control streams, an advanced solution is required to accommodate traffic classes and multiple streams, but there are no architectural specifics and the result. [5]

The eAVB supports time sensitive and low latency networks that guarantee Quality of Service (QoS). To support time sensitive applications in asynchronous automotive networking system. eAVB expands the norm by three additional sub-standards. [2,9,10] The eAVB composed of a bridge, a talker, and a listener. The gPTP (generic Precision Time Protocol) message is sent out by the talker and listener. It is submitted periodically to make up for mistake in time; then, the listener uses Multiple MAC Registration Protocol (MMRP) and Multiple Virtual LAN (VLAN) Registration Protocol (MVRP) to register VLAN and own mac addresses. The talker uses Multiple Stream Reservation Protocol (MSRP) to reserve the necessary bandwidth. The talker advertises the message to reserve the services, and the bridge calculates the required output port resource. The output is forwarded to the listener.

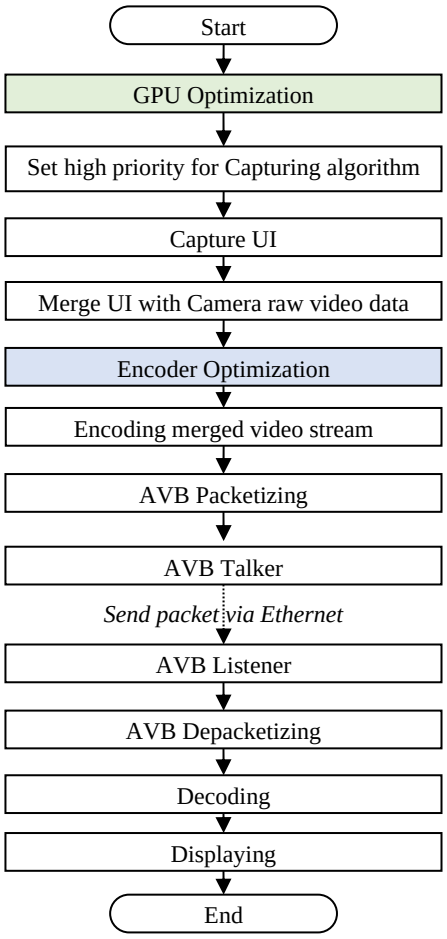


Fig. 2. Purposed RLMM flowchart.

3. Machine to Machine Network Design and Optimization of Latency

3.1 Machin to Machine Network Design

In this section, we present an architecture solution as shown in **Fig. 1**. It illustrates between components in automotive environment. The data stream is transmitted in the form of eAVB from the Ethernet camera, USB, and Antenna, and the data can be transmitted directly to each device, or to a device such as a head unit through an Ethernet switch. As depicted in **Fig. 1**, when data such as audio / video is sent directly to the head unit, the delay can be reduced. However, when other multimedia devices share the display with the head unit, a merge process of Camera's video and UI causes a big delay. In a vehicle, it is expected to increase the number of sharing a limited display with various multimedia because it's not practical to provide a display for every multimedia device unless it's a high-end car. As a result of actual experiments, merge the images by sharing the display as above, the H/U's display showed about twice as much delay as the image received directly from the camera.

3.2 Reduced Latency of Machine to Machine (RLMM)

3.2.1 Latency Computation

The latency computation of AVB *lQueue(m, q)* relies on the position, the reserved bandwidth of the queue *q*, and the message m on the traffic class. Consider *Q* the collection of AVB queues to route over *m*. (e.g., see Eq. 1, 2, 3, 4).

$$lQueue(m,\ q) = d_s(m,\ q),\ if\ q = q_s \quad (1)$$
$$lQueue(m,\ q) = d_0(m,\ q),\ if\ q = Q_0 \quad (2)$$
$$lQueue(m,\ q) = d_n(m,\ q),\ if\ q = Q_n \quad (3)$$

Consider q_s the queue with the lowest bandwidth required for the traffic class. When queues have the same lowest allocated traffic bandwidth, q_s is the largest of these. The total delay in shaping is applied at q_s. Where, Q_0 is the set of AVB queues that route the message m before queue q_s. Q_n is the set of AVB queues, the message m is routed over after q_s. It retains this as follows

$$Q = Q_0 \cup \{q_s\} \cup Q_n \text{ with}$$
$$Q_0 \cap \{q_s\} = Q_0 \cap Q_n = \{q_s\} \cap Q_n = \varnothing \quad (4)$$

3.2.2 RLMM

Through actual experiments between vehicle devices, we propose an optimized design. **Fig. 2** shows a simplified representation of the system. In the embedded multi-core ARM environment, the test results show that the CPU usage increases more than 65% when the AVB talker is run. In this situation, increasing the usage of CPU was considered undesirable, the experiment result shows that optimizes GPU can reduce the latency. It is important that the GPU acceleration meets the requirements of automotive applications while reducing resource usage of UI capture and audio / video encoding. We adopted GPU acceleration algorithm with minimal inrease in CPU usage. [11] By accelerating and optimizing GPU, it dramatically decreases the time of capturing UI. In addition, we used a SoC

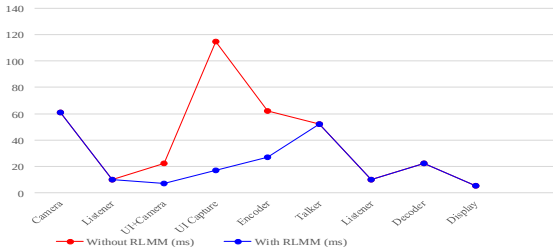


Fig. 3 Comparison between with RLMM and without RLMM

with optimized eAVB configuration to implement the design in Encoder. The purposed algorithm allows the delay reduction of all system aspects. There are many options for the encoding method, in consideration of the compression rate and resolution of the image according to the specifications requested by the automotive Original Equipment Manufacturer (OEM), it was confirmed that setting the specific coefficient in agreement with the vendor can reduce the delay. The selected Speed Dependent Motion Estimation (SDME) algorithm of encoder setting could decrease computational processing. [12] Section 4 explains better outcomes.

4. Experimental Results

To explore the feasibility of using Ethernet AVB, we developed a device model for comparing, validating, and evaluating Ethernet AVB's real-time performance to the current Ethernet standard. [2,3] **Fig. 1**, and **Fig. 2** visualizes the components and the test device model layout that consider the appropriate hardware and software components for the communication. The network consists of one Ethernet switch, and a variety of computers. The switch and devices are linked through a full duplex link. The experiment findings are presented in **Table 1**. Looking at the difference between without RLMM and with RLMM shown in **Fig. 3**, the merge process of UI and camera images took 22ms, UI capture 115ms, Video encoder 62ms, and AVB talker 52ms respectively. It showed the same level except for the latency of the multimedia device. According to our reliable experimental results, we obtained 7ms for UI and camera video merge video, 17ms for UI capture, 27ms for video encoder, and 7ms for AVB talker. Significant improvement of -15ms (68%) in UI and camera video merge,

Table 1. The latency of end-to-end components.

Device	Component	Without RLMM *(ms)*	Purposed RLMM *(ms)*	Reduced Latency *Δ (ms/%)*
Ethernet Camera		61	61	-
Ethernet Switch		1.5	1.5	-
Multi-media Device	Listener	10	10	-
	UI+Camera	22	7	-15 (68%)
	UI Capture	115	17	-98 *(85%)*
	Encoder	62	27	-35 *(56%)*
	Talker	52	52	-
Ethernet Switch		1.5	1.5	-
H/U	Listener	10	10	-
	Decoder	22	22	-
Display		5	5	-
Total Latency		362	169	-148 (41%)

-98ms (85%) in UI capture, -35ms (56%) in video encoder optimization The end-to-end processing time with the same payload is significantly less than the initial version. as described in **Table 1**.

Overall, The proposed RLMM method achieved about 41% performance improvement in total system.

4. Conclusions

In this paper, we have presented and investigated the latency of AVB network and suggested the optimized design of embedded system. The focus of this investigation lay on the influence of reducing delay between components. The proposed RLMM on realistic cases illustrated that there was a delay reduction about 41%. It is expected that the suggested method for the actual automotive environment can greatly shorten the time period in the design and development process. We not only presented a proof of latency analysis for AVB, but the alternative solution on practical test cases has also been tested. It is expected that the optimization method for the actual automotive environment can greatly shorten the time period in the design and development process. In-vehicle multimedia cameras generally use 2M (1920x1080), and this paper was also experimented with 2M standards. The algorithm proposed in this paper is expected to make a greater contribution when images of 8M (3840x2160) or more are introduced into the vehicle in the future. Therefore, it should be paid attention to it in design of future transmission with low latency for automotive.

Acknowledgement

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT).
(No. 2018-0-00824, Development of next generation image container standard for technology and structuring of relationship between 360-degree images)

References

[1] IEEE Standard for Local and metropolitan area networks, “Audio Video Bridging Systems,” in IEEE Std 802.1BA-2011, vol., no., pp.1-45, 30 Sept. 2011.
[2] IEEE Standard for Local and Metropolitan Area Networks, “Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks,” in IEEE Std 802.1AS-2011, vol., no., pp.1-292, 30 March 2011.
[3] J. Eveleens, "Ethernet AVB Overview and Status," SMPTE 2014 Annual Technical Conference & Exhibition, Hollywood, CA, USA, 2014, pp. 1-11.
[4] C. Herber, A. Saeed and A. Herkersdorf, "Design and Evaluation of a Low-Latency AVB Ethernet Endpoint Based on ARM SoC," 2015 IEEE 17th International Conference on High Performance Computing and Communications, NY, 2015, pp. 1128-1134.
[5] T. Nolte, H. Hansson and L. L. Bello, "Automotive communications-past, current and future," 2005 IEEE Conference on Emerging Technologies and Factory Automation, Catania, 2005, pp. 8 pp.-992.
[6] R. Queck, "Analysis of Ethernet AVB for automotive networks using Network Calculus," ICVES, Istanbul 2012, pp.61-67.
[7] X. Liu, Z. Nie, D. Li and H. Yu, "Design of An Improved Ethernet AVB Model for Real-time Communication in In-Vehicle Network," 2019 IEEE 3rd ITNEC, Chengdu, China, 2019, pp. 6-10.
[8] D. An, "Design and implementation of the car multimedia data transmission system using Ethernet AVB," 2016 Eighth ICUFN, Vienna, 2016, pp. 481-483.
[9] Virtual Bridged Local Area Networks – Amendment 9: Stream Reservation Protocol (SRP), IEEE P802.1Qat/D6.1 edition, June 2010.
[10] Virtual Bridged Local Area Networks – Amendment 12: Forwarding and Queuing Enhancements for Time-Sensitive Streams, IEEE P802.1Qav/D7.0 edition, Oct. 2009.
[11] J. M. Ready and C. N. Taylor, "GPU Acceleration of Real-time Feature Based Algorithms," 2007 IEEE Workshop on Motion and Video Computing (WMVC'07), Austin, TX, USA, 2007, pp. 8-8, doi: 10.1109/WMVC.2007.17.
[12] Byoungman An, Youngseop Kim and O. Kwon, "Low-Complexity Motion Estimation for H.264/AVC Through Perceptual Video Coding," KSII Transactions on Internet and Information Systems, vol. 5, no. 8, pp. 1444-1456, 2011.

A Survey on Deep Learning Approach in ECG Signals for Cardiac Arrhythmias Classification

Beanbonyka Rim[1], Jun Ma[1], Lyudmila Dmitrievna Khan[1] and Min Hong[2]
[1]Department of Software Convergence, Soonchunhyang University
Asan, South Korea
[2]Department of Computer Software Engineering, Soonchunhyang University
Asan, South Korea
[e-mail: rim.beanbonyka@gmail.com, ringring369@gmail.com, dmitrievna.km@gmail.com
minhong71@gmail.com]
*Corresponding author: Min Hong

Abstract

1D data detection and classification are currently the most potential for Deep Learning application. One of 1D data, the Electrocardiogram (ECG) data is used to diagnose cardiovascular diseases (CADs). Recently, ECG data has been fed into Deep Learning method to do an automatic CADs classification. In this paper, we present a comprehensive study on taxonomy, architecture, processes and key challenges of Deep Learning approach in ECG signals for CADs classification. We reviewed 11 papers from the recent publications within year 2020. Our study provides an evidence-based support for critically understanding the core processing stages such as data acquisition, pre-processing, feature extraction, and processing. We also highlight the supporting stages such as prediction and evaluation metrics.

***Keywords*:** Electrocardiogram data, 1D data classification, Deep Learning, cardiovascular diseases

1. Introduction

Deep Learning (DL) approach has proved a significiant and reliable performance on an automatic medical imaging analysis, recently. Manual feature extraction (text data which extracted by expert or cardiologist) is always and potentially a burden in both resource and time-comsuming for classification task. [1]. Therefore, in this paper, we present a comprehensive and comparative study on taxonomy, architecture, processing and key challenges of DL approach in ECG signals for CADs classification. We collected 11 published papers via search engine PubMed with keywords combining of "ecg deep learning" which only published in year 2020 [2].

2. ECG Analysis Architecture

We describes typical stages of DL approach in ECG analysis architecture for CADs classification by the 11 research works [3-13].

2.1 Data Acquisition

The ECGs are available to be captured from a 1-lead to 12-leads recording devices such as ECG electrodes-amplifier, wireless sensors, and

This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Korean government (MSIT) (No. NRF-2019M3E5D1A02069073).

mobile or portable sensors. The 12-leads ECG records are I, II, III, aVR, aVL, aVF, V1, V2, V3, V4, V5 and V6 [7]. The well-known public datasets are such as MIT-BIH Arrhythmia, INCART, Fantasia and CinC of Physionet database and China Physiological Signal Challenge (CPSC).

2.2 Pre-processing

The most common method for removing noise such as discrete wavelet transforms (DWT) in which is used by the works of [3,4] and for normalizing such as z-score method, in which is used by the works of [3].

2.3 Feature Extraction

The work of [3] segmented the signals to 2s for short beat and 5s for long beat. Other works defined fixed length such as the works of [4,12] segmented the signals to 10s, the work of [10] segmented the signals to 60s. The work of [13] converted into 2D spectrogram images using short-time fourier transform (STFT).

2.4 Processing

Table 1 describes the DL models which were deployed for CADs in ECG signals.

Table 1. The summaries of DL model deployment

Papers	DL model
Wang et al. [7]	CNN
Ribeiro et al. [8]	CNN
Van de Leur et al. [9]	37-layer Resnet
Romdhane et al. [11]	CNN
Wang et al. [4]	GoogLeNet, Resnet, SeNet+ Feature fusion LightGBM Voting
Ullah et al. [13]	2D-CNN
Chang et al. [6]	LSTM
Butun et al. [3]	CNN Capsule net Decorder
Chen et al. [5]	CNN GRU Attention Dense
Zhang et al. [10]	VGG STA bi-GRU
Yildirim et al. [12]	CNN LSTM

2.5 Prediction

In 11 research works, there is one work [3] who conducted a binary classification, i.e. a normal ECGs or an abnormal ECGs (CADs). The others conducted multi classified categories from 4 to 34 classes.

2.6 Evaluation Metrics

The typical evaluation metrics are accuracy, sensitivity, specificity, precision, recall, F1 score, AUC, listed details in **Table 2**.

Table 2. Performance results

Papers	Result
Butun et al. [3]	2s: Sen: 99.70%, Spe: 98.10%, Acc: 99.44% 5s: Sen: 98.75%, Spe: 97.97%, Acc: 98.62%
Wang et al. [4]	F1: 0.9238 Precision: 0.9372 Recall: 0.9107
Chen et al. [5]	F1: 0.84
Chang et al. [6]	Acc: 0.982-1.0 AUC: 0.987-1.0 Precision: 0.692-1.0 Recall: 0.625-1.0 Test set: Acc: 90%
Wang et al. [7]	12-lead: F1: 82.8%, Pre: 83.8%, Recall: 82.2% Single-lead: F1: 84.1%, Pre: 85.6%, Recall: 82.9%
Ribeiro et al. [8]	F1: > 88% Spe: > 99% Pre: > 83% Recall: > 76%
Van de Leur et al. [9]	Sen: 0.8025 Spe: 0.9225
Zhang et al.[10]	F1: 0.835
Romdhane et al. [11]	overall accuracy: 98.41% overall F1-score: 98.38% overall precision: 98.37% overall recall: 98.41%
Yildirim et al. [12]	7 reduced rhythms: Acc: 92.24% 4 merged rhythms: Acc: 96.13%
Ullah et al. [13]	Acc: 99.11%

4. Conclusions

In this paper, we present a comprehensive and comparative study on taxonomy, architecture, processing and key challenges of DL approach in ECG signals for CADs classification. The 11 published papers which classify ECG signals

into binary or multi classes of abnormalities of cardiovascular, were deeply reviewed, analyzed and taxonomized in each processing stage.

References

[1] B. Rim, J. Kim, and M. Hong, "Deep Learning in Physiological Signal Data: A Survey", *Sensors*, Vol. 20, No. 4, p. 969, 2020. https://doi.org/10.3390/s20040969

[2] PubMed. Available online: https://www.ncbi.nlm.nih.gov/pubmed/ (accessed on 31 September 2020).

[3] E. Butun, O. Yildirim, M. Talo, R.S. Tan, and U.R. Acharya, "1D-CADCapsNet: One dimensional deep capsule networks for coronary artery disease detection using ECG signals", *Physica Medica*, Vol. 70, pp.39-48, 2020. https://doi.org/10.1016/j.ejmp.2020.01.007

[4] D. Wang, Q. Meng, D. Chen, H. Zhang, and L. Xu, "Automatic Detection of Arrhythmia Based on Multi-Resolution Representation of ECG Signal", *Sensors*, Vol. 20, No. 6, p.1579, 2020. https://doi.org/10.3390/s20061579

[5] T.M. Chen, C.H. Huang, E.S. Shih, Y.F. Hu, and M.J. Hwang, "Detection and Classification of Cardiac Arrhythmias by a Challenge-Best Deep Learning Neural Network Model", *Iscience*, Vol. 23, No. 3, p.100886, 2020. https://doi.org/10.1016/j.isci.2020.100886

[6] K.C. Chang, P.H. Hsieh, M.Y. Wu, Y.C. Wang, J.Y. Chen, F.J. Tsai, E.S. Shih, M.J. Hwang, and T.C. Huang, "Usefulness of Machine-Learning-Based Detection and Classification of Cardiac Arrhythmias with 12-Lead Electrocardiograms", *Canadian Journal of Cardiology*, In press, available online, 05 March 2020. https://doi.org/10.1016/j.cjca.2020.02.096

[7] R. Wang, J. Fan, and Y. Li, "Deep Multi-Scale Fusion Neural Network for Multi-Class Arrhythmia Detection", *IEEE journal of biomedical and health informatics*, Vol. 24, No. 9, pp.2461-2472, 2020. https://doi.org/10.1109/JBHI.2020.2981526

[8] A.H. Ribeiro, M.H. Ribeiro, G.M. Paixão, D.M. Oliveira, P.R. Gomes, J.A. Canazart, M.P. Ferreira, C.R. Andersson, P.W. Macfarlane, M. Wagner Jr, and T.B. Schön, "Automatic diagnosis of the 12-lead ECG using a deep neural network", *Nature communications*, Vol. 11, No. 1, pp.1-9, 2020. https://doi.org/10.1038/s41467-020-15432-4

[9] R.R. van de Leur, L.J. Blom, E. Gavves, I.E. Hof, J.F. van der Heijden, N.C. Clappers, P.A. Doevendans, R.J. Hassink, and R. van Es, "Automatic Triage of 12- Lead ECGs Using Deep Convolutional Neural Networks", *Journal of the American Heart Association*, Vol. 9, No. 10, p.e015138, 2020. https://doi.org/10.1161/JAHA.119.015138

[10] J. Zhang, A. Liu, M. Gao, X. Chen, X. Zhang, and X. Chen, "ECG-based multi-class arrhythmia detection using spatio-temporal attention-based convolutional recurrent neural network", *Artificial Intelligence in Medicine*, Vol. 106, p.101856, 2020. https://doi.org/10.1016/j.artmed.2020.101856

[11] T.F. Romdhane, and M.A. Pr, "Electrocardiogram heartbeat classification based on a deep convolutional neural network and focal loss", *Computers in Biology and Medicine*, Vol. 123, p.103866, 2020. https://doi.org/10.1016/j.compbiomed.2020.103866

[12] O. Yildirim, M. Talo, E.J. Ciaccio, R. San Tan, and U.R. Acharya, "Accurate deep neural network model to detect cardiac arrhythmia on more than 10,000 individual subject ECG records", *Computer methods and programs in biomedicine*, Vol. 197, p.105740, 2020. https://doi.org/10.1016/j.cmpb.2020.105740

[13] A. Ullah, S.M. Anwar, M. Bilal, and R.M. Mehmood, "Classification of Arrhythmia by Using Deep Learning with 2-D ECG Spectral Image Representation", *Remote Sensing*, Vol 12, No. 10, p.1685, 2020. https://doi.org/10.3390/rs12101685

A Design of Multi-Condition Emotional Speech Synthesizer

Sung-Woo Byun and Seok-Pil Lee
Department of Computer Science, The graduate school, SangMyung University
Seoul, South Korea
[e-mail: 123234566@naver.com, esprit@smu.ac.kr]
*Corresponding author: Seok-Pil Lee

Abstract

Recently, researchers proposed text-to-speech models based on deep-learning whose results were superior to the previous approaches. However, since these models only referred speaker's speaking style of reference speeches, those were difficult to reflect speaker's emotions. This paper proposes an emotional speech synthesizer by embedding not only speaking styles but also emotion styles. We extend speaker embedding to multi-condition embedding by adding emotion embedding in Tacotron so that the synthesizer can generate emotional speech. Evaluation test results show the superiority of the proposed algorithm to the previous methods in terms of emotional expressiveness.

***Keywords*:** Volumetric medical image, image segmentation, surface reconstruction, Augmented Reality

1. Introduction

Text-to-speech system (TTS) synthesizes and output human-like speech signals so that linguistic, prosodic, and semantic information can be naturally delivered. Recently, due to advances in deep learning technologies, researchers have developed text-to-speech models based on deep learning, which have produced results superior to those of previous approaches. This work led to investigations into the issue of understanding the contextual meaning of synthesized speech [1]. However, since these models only took into account a speaker's style of presenting reference speeches, it was difficult to reflect the speaker's emotions.

Deep learning-based TTS systems such as Tacotron produce expressive speech by adding embedding vectors that implicitly provide prosody-related latent features [2]. However, because those systems only mimic the generic speaking style of reference audio, it is difficult to assign user-defined emotion types to synthesized speech [1]. Lee et al. [3] input emotion labels to the decoder of Tacotron by concatenating the labels with the output of the pre-net. Training using this approach can make the model reflect the emotional state of a speaker. Even though the method showed feasibility when attaching emotions to synthesized speech, when the amount of data is small, there is still a problem with limited emotional representation in the speech.

In this paper we propose an emotional speech synthesizer constructed by embedding not only speaking styles but also emotion styles. We extend speaker embedding to multi-condition embedding by adding emotion embedding in Tacotron, so that the synthesizer can generate emotional speech. The model can be trained on the emotional style of speakers, and the use of multiple speakers means a lack of emotional state in one speaker's speech data can be complemented with data from other speakers. This approach facilitates training the attention of

the model, by training using an emotional speech dataset with a large amount of speech data. Evaluation of the results showed the superiority of the proposed model to a previous model, in terms of emotional expressiveness.

This paper is organized as follows. Section 2 introduces researches on existing speech synthesizer technology. Section 3 discusses multi-condition emotional speech synthesizer, which are the key parts of the proposed method. Section 4 presents the experiment description and results, and then discusses and concludes with Sections 5.

2. Related works

2.1 Tacotron

Among the TTS models [4, 5], Wang et. al [2] proposed an end-to-end TTS model, referred as Tacotron, that can be trained from scratch on <text, audio> pairs. Tacotron is a sequence-to-sequence (seq-to-seq) model with an attention mechanism. Tacotron has many advantages compared with other TTS models such as WaveNet [4] and Deep Voice [5]. The speech signal generated by Tracotron achieved competitive mean opinion scores (MOSs).

2.2 Emotional Speech synthesizer

Lee et al. proposed a modified Tacotron as an emotional speech synthesizer. The system takes a character sequence and a desired emotion as input, and generates the corresponding wave signal. The model was implemented by injecting a learned emotion embedding, *e*, and demonstrated the feasibility of adding emotions to synthesized speech.

3. Multi-condition emotional speech synthesizer

The objective of the proposed method was to train a TTS system using speech data from multiple speakers, so that the emotional states of different speakers complement each other. Unlike previous studies, we extended a single embedding vector, such as a speaker embedding vector [6] or an emotion embedding vector, so that the model can reference multiple conditions. Each of the multi-condition embeddings were initialized randomly with a uniform distribution over [−0.1, 0.1] and the network was trained using backpropagation.

$$\begin{aligned} z &= concat(s, e) \\ Multicondition &= \tanh(w_{f,n} * z) \end{aligned} \tag{1}$$

where *s* is a speaker embedding and *e* is an emotion embedding. The multi-condition embeddings are fed into the model when training.

The embeddings are combined with a CBHG module (Convolution Bank + Highway + bi-GRU), a Long Short-Term Memory (LSTM) encoder, attention recurrent neural networks (RNNs), and a decoder RNN equivalent to Deep Voice 2 [6]. The whole structure is shown in **Fig. 1**.

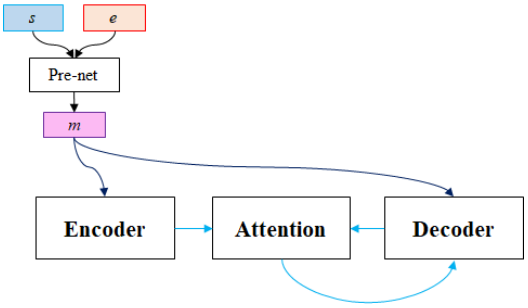


Fig. 1. The whole structure of Multi-condition emotional speech synthesizer

4. Materials and experiment

4.1 Dataset

In this research, A Korean emotion database was constructed. The database was recorded from Korean scripts from dramas and movies, and the utterances consist of conversations of males and females. We defined four emotions: anger, happiness, neutrality, and sadness.

All data were recorded with four professional actors, two female and two male (**Fig. 2**), in a professional studio, to ensure good sound quality without any noise. All speech emotion data was recorded at 48kHz and downsampled with a 24kHz downsampling rate in PCM signed 16-bit format.

Fig. 2. The two actors during recording

4.2 Experiment

To complement the small size of the constructed dataset, we trained our emotional speech synthesizer by adding another dataset of eight hours. We used speech data from two speakers in the TTS dataset and four speakers in the emotional speech dataset.

To demonstrate the clarity and expressiveness of the proposed model, a five-scale MOS test was carried out with five subjects. Each subject listened to each sample and was asked to decide the clarity on a scale of 1 to 5, according to whether they could understand the words in the sample. They were also asked to provide an emotion score from 1 to 5. If a subject felt that the emotion was clearly evident in the speech data, they gave a score of 5. However, if a subject felt as if the speech data did not include emotion, they give a score of 1.

Table 1. 5-scale mean opinion score evaluation.

	Clarity	Emotion Score
Prpposed	4.07±0.81	4.11±1.32
[3]	3.52±1.08	3.77±1.10
Real data	4.95±0.21	4.8 ±0.5

As shown in **Table 1**, the proposed model achieved a clarity of 4.07, which outperforms the previous emotional speech synthesizer with 3.52. The proposed model achieved better emotion score of 4.11 than the other model, which had a score of 3.77.

5. Conclusions

This paper describes an emotional speech synthesizer constructed by extending speaker embedding to multi-condition embedding by adding emotion embedding. The model was trained using the emotional style of the speakers, and a lack of an emotion in one speaker was complemented by data from other speakers. To facilitate training the attention of the model, we trained the model using both the emotional speech dataset and a large amount speech data for TTS. A five-scale MOS test was carried out using five subjects. The experimental results showed that the proposed model outperformed the previous model in terms of emotional expressiveness.

Acknowledgement

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. NRF-2019R1F1A1050052)

References

[1] S. Y Um, S. S Oh, K. G Byun etal., “Emotional Speech Synthesis with Rich and Granularized Control,” *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), 2020.

[2] Y. Wang, RJ Skerry-Ryan, D. Stanton etal., "Tacotron: Towards End-to-End Speech Synthesis", *arXiv preprint arXiv:1703.10135*, 2017.

[3] Y. G Lee, A. Rabiee, S. Y Lee., " Emotional End-to-End Neural Speech Synthesizer", *arXiv preprint arXiv:1711.05447*, 2017.

[4] A. Oord, S. Dieleman, H. Zen, K. Simonyan, etal., “Wavenet: A generative model for raw audio,” *arXiv preprint arXiv:1609.03499, 2016.*

[5] S. Ö. Arık, M. Chrzanowski, A. Coates, etal., "Deep voice: Real-time neural text-to-speech," *Proceedings of the 34th International Conference on Machine Learning*, 2017.

[6] [4] S. Arik, G. Diamos, A. Gibiansky, etal., “Deep Voice 2: Multi-Speaker Neural Text-to-Speech,” *arXiv preprint arXiv:1705.08947*, 2017.

KSII The 12th International Conference on Internet (ICONI) 2020.

A CNN-based Place Recognition with Attention Method in Broadcasting Video

Giseok Choe and Jongho Nang
Department of Computer Science and Engineering, Sogang University
Seoul, South Korea
gschoe@sogang.ac.kr, jhnang@sogang.ac.kr
*Corresponding author: Jongho Nang

Abstract

This thesis proposes a place classifier in broadcast video. In the proposed method, only a fixed number of frames were sampled to be useful for scene recognition with a sampling method using binary classifier and extracted feature. Its strength was adjusted through the analysis of features using the attention method to determine whether there was noise on frame and fused to create a feature of the scene, the fusion method used pooling and clustering. Scene is recognized using fully connected method classifier to use a feature of the scene. For scene recognition, use classifier the fused features. Use image-based dataset for training and validation the proposed methods. Take 2,464 scenes in 8 movies and 10 dramas from real video data for testing of performance fusion features and classifier. As a result, it is show that Top-1 35% and Top-5 66% accuracy about scene recognition.

***Keywords*:** Video, Place Recognition, Attention, CNN, Deep Learning

1. Introduction

The development of smart phones and communication technology has exploded the supply and demand for video. Therefore, many studies are being conducted to automatically analyze video data. In this video analysis, place is important metadata information. So, there have been many studies to recognize a place from a single image. However, the video of broadcast content has different characteristics from a single image. Since the video content has a characteristic that the location does not change in units of frames, there is an advantage in that a lot of information can be simultaneously inputted in units of scenes. Since most of the frames are centered on people, it is difficult to infer the background by looking at a single frame. Existing studies have not considered the characteristics of broadcast contents. In this paper, we propose a background recognition system suitable for broadcasting contents by extracting a number of frames from a video by scene unit and selecting frames that are advantageous for location recognition.

Through the analysis of the broadcasting contents, a place category suitable for the analysis of the broadcasting video was selected. In addition, a binary classifier-based frame sampling method is proposed to extract frames suitable for place recognition. The deep feature set was extracted for the extracted frames, and fused using pooling and NetVLAD[1] methods. In this process, the attention method was used to emphasize a frame that is better for place recognition for each frame feature.

This work was partly supported by Institute of Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (2019-0-01730. Development of a Video Fingerprint and Fast Retrieval Method for Preventing the Wide-spreading of Sexual Crime Videos)

2. Related Research

Place recognition was mainly based on image information. Methods to improve Place recognition performance include dividing an image into multiple patches[2], or adding additional information such as objects[3]. Recently, these studies have improved greatly with the development of CNN-based deep learning. Deep features using network structures such as VGG[4] or ResNet[5] showed superior results compared to existing hand-crafted features. In addition, as datasets such as Places365[6] are released, their research is also being actively conducted. Methods of using multiple images obtained from moving pictures include a method of recognizing a place by accumulating the features of previous frames[7] or a method such as RNN[8]. Recently, attempts have been made to increase the performance by utilizing the attention method[9].

In the case of YUPenn/YUP++[10] or Maryland[11], the focus is on solving the problem of location recognition for a single short shot within 5 seconds. Coview' 18[12] can be used to solve the problem of location recognition for scenes where several relatively long shots are clustered, but there are only 29 labels in total, and the composition is also difficult to use in analyzing broadcast contents. I had it and it was not suitable.

3. System Design

In order to recognize a place in a scene unit, a structure like the **Fig. 1** was proposed, and it can be divided into four types: frame sampling method, feature extraction method, fusion method, and classification method.

3.1 Frame Sampling Method

Since there are many depictions focused on people in broadcast content, frames that cannot be identified often appear. Therefore, it is important to select an appropriate frame to determine the place. To solve this problem, a binary classifier as shown in Figure 2 was used. This classifier changed the fully-connected layer of ResNet-50[5] network trained with Places365[6] into a binary classifier and performed transfer learning. Through this, the upper 32 frames can be extracted by arranging the frames in an order that is easy to detect the place.

3.2 Feature Extraction Method

As a feature vector for each frame, a 2,048-dimensional vector, which is the result of the pre-coupling layer, was used in the ResNet-50[4] network learned with Places365[6]. For 32 frames, feature vectors for each frame are extracted, and a vector of 32×2,048 dimensions is extracted and used by using this.

A feature vector for each scene is created through a fusion process through the feature vectors for each frame with adjusted weights. In this paper, two methods were used to simplify a vector of 32×2,048 dimensions to a vector of 2,048 dimensions. The first is vector generation through average pooling. The second is the feature fusion method using the NetVLAD[1] method.

3.3 Fusion Method

The attention method as shown in Fig. 2 was used to select the main information among the information spanning several frames. For the extracted features of 32 frames, a

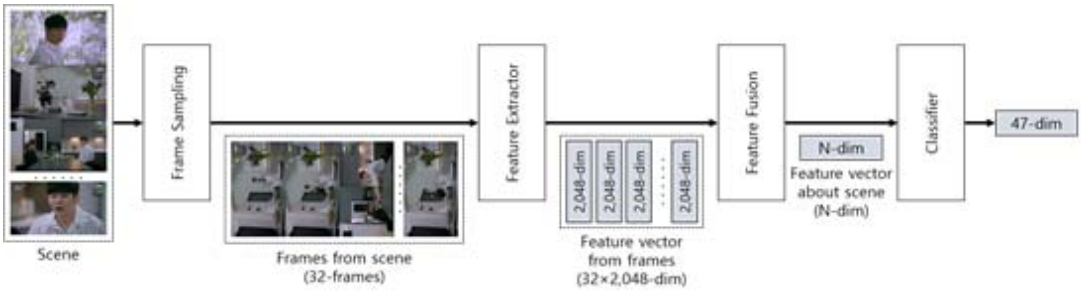


Fig. 1. Architecture of scene-level place recognition in broadcasting video

32-dimensional attention map is constructed by analyzing features for each feature using a CNN-based temporal attention layer. The attention map adjusts the weight for each frame.

3.4 Classification Method

Place classification was performed using a fully-connected layer, which is often used in the existing deep learning method.

4. Experiments

4.1 Environments

In the experiment of this paper, we implemented using Intel Core i7-7700K 4.20GHz as CPU, NVIDIA GeForce GTX 1080 Ti hardware as GPU, Ubuntu 16.04 as operating system, and 0.4.1 version of PyTorch deep learning framework. .

4.2 Category Definition

Places365[6] divided the place category so finely that it did not fit the broadcast content. Therefore, it was necessary to define a place category for place recognition of broadcast content. Broadcast videos that aired in Korea were collected. As for the collected broadcasting contents, 8 movies and 10 dramas were used as shown in Table 1. A total of 51 place categories were extracted by comparing the place categories and relationships of Places365[6]. Among them, 47 categories were finally defined by removing 4 labels specialized for a specific country or genre (palace, ondol, rooftop, executive office). The defined categories are shown in **Table 2**.

Table 1. List of broadcast videos used in the dataset

Type	*Name*
Movie	A Taxi Driver, Covertly Gloriously, Love Forecast, New World, The Last Blossom, Take Off, Confidential Assignment, Steel Rain
Drama	The Last Empress (E33~E34), My Only One (E71~E72), Clean With Passion For Now (E11), The Crowned Clown (E04), My Strange Hero (E19~E20), Miss Hammurabi (E09), My ID is Gangnam Beauty (E09), Encounter (E09), My Mister (E09), What's Wrong with Secretary Kim (E09)

Table 2. List of place categories defined by referring to broadcast video and Places365 dataset

Place47 Categories
Inside a car, roadside/walkway, restaurant, living room, office, warehouse/basement, classroom, road/lane, palace, bedroom, ondol room, hospital, rooftop, kitchen, executive room, construction site, building exterior, corridor, stage/stage, conference room, Pub/bar, factory, indoor/underground parking lot, mountain/field, outdoor parking lot, cafe/coffee shop, shop, cinema/auditorium, park/square, playground/amusement park, stairs/escalator, river/riverside, public transport, lobby , Countryside, Sea/Beach, Yard/Garden, Court, Playground, Toilet, Tunnel/Cave, Entrance, Pier/Port, Studio, Prison, Laboratory, Elevator, Airport, Library, Gymnasium, Exhibition Facility

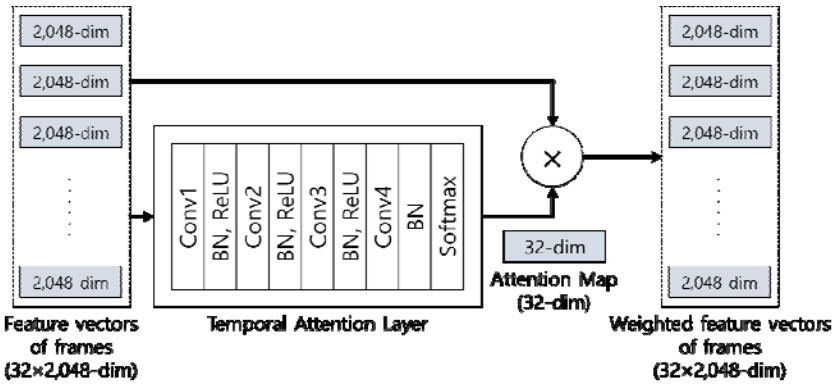


Fig. 2. Structure of the attention method

4.3 Dataset

First, the existing Places365[5] dataset was converted to fit the 47 defined categories. The data converted in Place365 was classified as an image that showed the place well. In addition, 120,000 images that do not reveal a place were collected and added.

In the movie and drama used in **Table 1**, 24 episodes of the American drama "Friends Season 1" were added, and 47 locations were labeled based on the scene. As a result, 1,892 scenes for movies, 572 scenes for dramas, and 336 scenes for Friends were configured as evaluation datasets.

4.3 Experiment Results

Table 3. Accuracy of binary classifier for frame sampling

Dataset	*Accuracy(%)*
Training	75.66
Validation	76.18

First, in order to experiment on the sampling performance, images were mixed with positive, negative, and noise sets, and learned and verified. The results are shown in **Table 3**.

Experiment by forcibly changing the number of positive images. As a result of the experiment, the presence and performance of the attention method were confirmed. As shown in Fig. 3, it can be seen that the attention weight increases to compensate for the decrease in the number of positive images, and the overall performance increases by 1-5% compared to when the attention is not used.

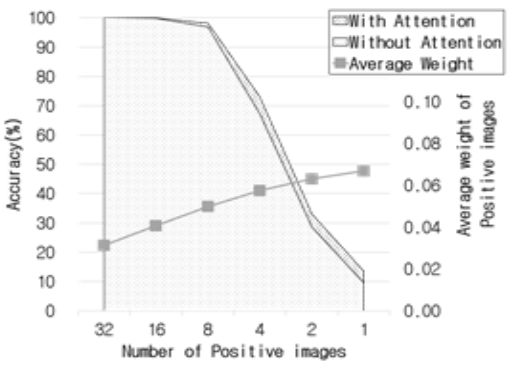


Fig. 3. Accuracy and weight according to the number of positive images

Next, the number of positive images was fixed to 4 and the ratio of the negative image and the noise image was changed to confirm the performance. As a result, as shown in **Fig. 4**, in the case of the negative image, the complement effect due to attention was less than that of the noise image.

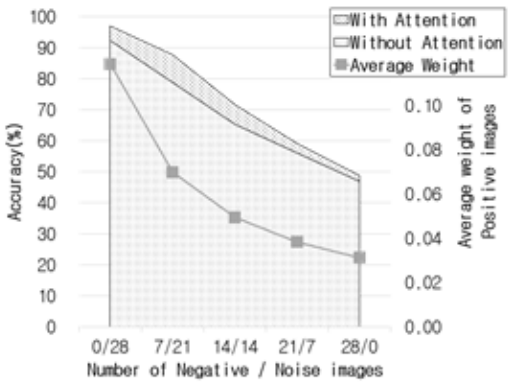


Fig. 4. Accuracy and weight according to the number of negative and noise images

Fig. 5 shows the results of comparing the performance to confirm the performance according to the fusion method. It can be seen that the netvlad method shows slightly better performance than the average pooling method.

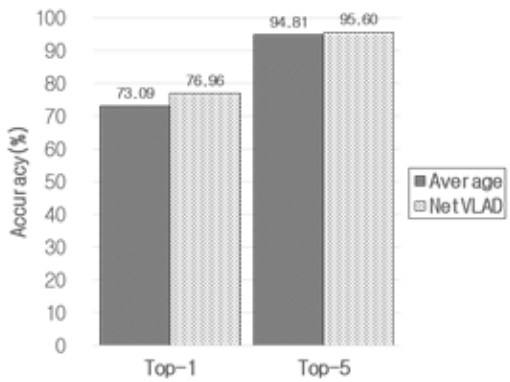


Fig. 5. Performance comparison according to the fusion method

Fig. 6 shows the combined performance of attention and fusion methods. The average pooling method applying attention showed the best performance, with 35% of Top-1 and 66% of Top-5.

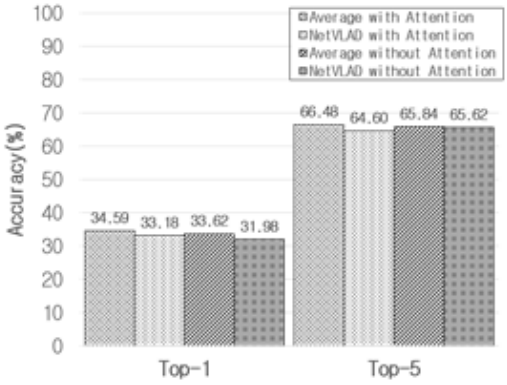


Fig. 6. Performance comparison according to attention and fusion method

5. Conclusions

In this paper, a method for data definition and system design for performing scene-based place recognition for broadcast content is proposed. 47 place labels were reconstructed in consideration of the location recognition characteristics of broadcast content. When multiple frames are input, a binary classifier-based sampling structure is proposed to select a frame with a good location to improve the performance of the entire system. Also, for the sampled frames, features were extracted using the ResNet-50[5] network learned with Places365[6]. Feature vectors suitable for place recognition were created using the attention method and the fusion method.

A new dataset was constructed to learn and verify the proposed method. The dataset was composed by mixing the Places365 dataset[6] and the collected videos and images. As a result, it was possible to obtain performance of 35% in Top-1 criteria and 66% in Top-5 criteria for location recognition of broadcast videos.

In the future, we will proceed with further research on a method of integrating a network divided into several stages and a data set composition that is easier to verify.

References

[1] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla, and J. Sivic, “NetVLAD: CNN Architecture for Weakly Supervised Place Recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 5297-5307.

[2] A. Quattoni, and A. Torralba, "Recognizing indoor scenes." pp. 413-420.

[3] P. Espinace, T. Kollar, A. Soto, and N. Roy, "Indoor scene recognition through object detection." pp. 1406-1413.

[4] K. Simonyan, and A. Zisserman, “Very Deep Convolutional Networks for Large-scale Image Recognition,” *arXiv preprint arXiv:1409.1556*, 2014.

[5] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going Deeper with Convolutions,” in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 2015.

[6] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, “Places: A 10 Million Image Database for Scene Recognition,” *IEEE transactions on pattern analysis and machine intelligence*, 2017.

[7] J. Wu, H. I. Christensen, and J. M. Rehg, "Visual place categorization: Problem, dataset, and algorithm." pp. 4763-4770.

[8] Y. Huang, X. Cao, Q. Wang, B. Zhang, X. Zhen, and X. Li, “Long-Short-Term Features for Dynamic Scene Classification,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 4, pp. 1038-1047, 2018.

[9] J. Park, S. Jeon, S. Kim, J. Lee, S. Kim, and K. Sohn, "Learning to detect, associate, and recognize human actions and surrounding scenes in untrimmed videos." pp. 21-26.

[10] K. G. Derpanis, M. Lecce, K. Daniilidis, and R. P. Wildes, "Dynamic scene understanding: The role of orientation features in space and time in scene classification." pp. 1306-1313.

[11] N. Shroff, P. Turaga, and R. Chellappa, "Moving vistas: Exploiting motion for describing scenes." pp. 1911-1918.

[12] K. Sohn, M.-H. Yang, H. Byun, J. Lim, J. Hsu, S. Lin, E. Kim, and S. Kim, "Coview'18: The 1st workshop and challenge on comprehensive video understanding in the wild." pp. 2113-2115.

AR based realistic cloth simulation using spatial mapping

Jun Ma[1], Lyudmila Dmitrievna Khan[1], Va Hongly[1] and Min Hong[2]
[1]Department of Software Convergence, Soonchunhyang University
Asan, South Korea
[2]Department of Computer Software Engineering, Soonchunhyang University
Asan, South Korea
[e-mail: ringring369@gmail.com, dmitrievna.km@gmail.com, vahonglykhmer@gmail.com
minhong71@gmail.com]
*Corresponding author: Min Hong

Abstract

This paper is a Realistic cloth simulation program using spatial mapping in augmented reality, which allows the cloth object to operate realistically in augmented reality environment, causing a collision reaction with a mesh produced through the spatial mapping. However, the current spatial mapping consumes a lot of memory during the mesh collision, resulting in significant frame drop. Therefore, this paper presents the simplification algorithm of mesh collision and the frame rate comparison of the existing spatial mapping cloth simulation, and the new spatial mapping cloth simulation accordingly.

***Keywords*:** Augmented reality, Cloth simulation, Spatial mapping

1. Introduction

The cloth simulation has been widely applied to express realistic and immersive scenes in augmented reality. In general, conventional cloth simulations have focused on realism rather than real-time processing. In this case, it may take dozens of minutes to achieve the rendering by building a sophisticated cloth model and performing complex simulations with it. However, in case of real-time augmented reality, the simulation approach either applies GPU-based methods or reduces mesh resolution to increase performance[1].

In particular, efforts were made to achieve the advanced performance of collision detection, a major bottleneck in computer graphics. For faster collision detection, the sphere or capsule collider surrounding the complex model is first defined, then the mesh is simplified and the collision detection of the corresponding collider and cloth model is carried out[2]. Augmented reality (AR) content is a representative form of interactive content. AR content basically requires expensive calculation costs to search for targets in camera image sequences. Thus, in order to implement dynamic simulation between 3D objects that conform to image targets in AR, a highly efficient dynamic simulation method for real-time processing should be applied. To improve performance while maintaining rendering quality in real-time, deformable objects were simulated with a low-resolution mesh combined with high-resolution surface mesh used to visualize deformable bodies[3].

In this paper, a cloth simulation method including efficient collision detection and processing algorithm for AR content is proposed. In the proposed method, objects are defined using spatial mapping, and efficient collision handling between 3D objects and fabric models

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2019R1F1A1062752)

is used for rendering optimization. We develop it as a plug-in to Unity Engine, which is widely used in real-time game development. The performance improvement of the proposed method is presented through comparative experiments with existing models and applied to AR contents.

2. AR based realistic cloth simulation

This part describes the implementation of optimization algorithm for spatial mapping mesh and the implementation of cloth simulation that responds to that mesh.

2.1 Optimization of spatial mapping mash Algorithm

```
(1) SetThresholds(t1 ← min Y, t2 ← 2/4 Y, t3 ← 3/4 Y, t4 ← max Y)
(2) k-Mean4(thresholds, meshVertices)
(3)     forall meshVertices
(4)         calculate Euclidian Distance between a threshold t and a vertex v
(5)     Sort calculated distances
(6)     Include a vertex v into the cluster with the closest threshold t
(7)     Output clusters(c1, c2, c3, c4)
(8)UpdateT()
(9)     forall i from 0 to 4
(10)        newTi ← average height inside the ci
(11)do k-Mean4(), UpdateT() till (t - newT) > ε
(12)SetK()
(13)    forall i from 0 to kNumber
(14)        ki ← random v from c
(15)k-MeanSubclusters()
(16)    forall meshVertices in cluster c
(17)        calculate Euclidian Distance between vertex v and each threshold k
(18)    Sort calculated distances
(19)    Include a vertex v into the cluster with the closest threshold k
(20)    Output subclusters(S1,..., Sknumber)
(21)UpdateK()
(22)    forall i from 0 to kNumber
(23)        newKi ← average vertex position from subcluster Si
(24)do k-MeanSubclusters(), UpdateK() till (k - newK) > ε
(25)SetNewVertices()
(26)    newMeshVert ← k
```

1. Set threshold Y based on initial height.
2. Divide the mesh into four sections according to the threshold Y set.
3. After dividing, reset threshold Y to the average value within the interval.
4. Then use the new threshold T to divide the mesh back into four sections.
5. Repeat the reset and the division of the mesh, comparing the threshold T to the previous threshold t.
6. Then create a new threshold coordinate K. Sets a new threshold K based on the location of the mesh vertex.
7. Divide each section into smaller sections. The threshold K is then reset to the average value. Repeat reset and mesh division, comparing threshold K with last threshold k.
8. Proceed through the above process to simplify the mesh.

To verify the optimization of the mesh, we conducted the test using the Stanford bunny model from **Fig. 1**.

Fig. 1. Stanford bunny model.

The model is basically composed of 34,835 vertices and 69,666 faces, and we divided it into 100 subclusters to optimize. The results are shown in **Fig. 2**.

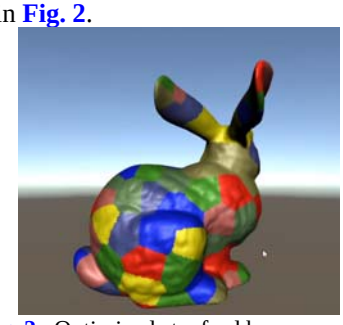


Fig. 2. Optimized stanford bunny model.

The collision of the cloth simulation of the model was then applied as shown in **Fig. 3**.

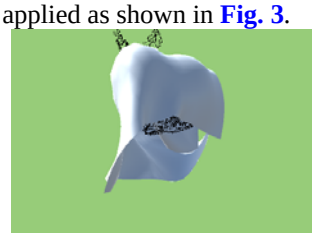


Fig. 3. Cloth simulation with stanford bunny model.

2.2 Realistic cloth simulation in AR Spatial mapping

We extracted the mesh model of the room using the spatial mapping technology of Vive pro and proceeded with optimization. It's shown in **Fig. 4**.

Fig. 4. Room model

After extracting the characteristics of the mash model, the process of leaving only the points to be used for mesh optimization was carried out as shown in **Fig. 5**.

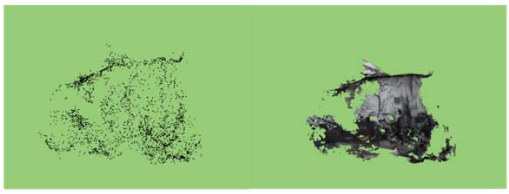


Fig. 5. Feature point extraction

The screens before and after mesh optimization are shown in **Fig.s 6** and **7**.

Fig. 6. Before spatial mapping optimization

Fig. 7. After spatial mapping optimization

Using this optimization algorithm, we applied cloth simulation as shown in **Fig. 8**.

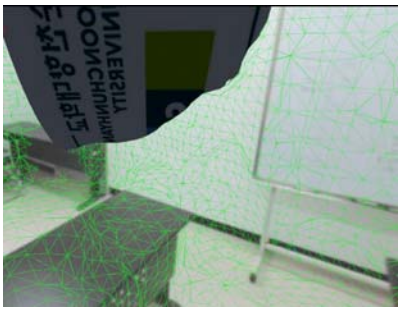


Fig. 8. Cloth model with spatial mapping

3. Conclusions

In this paper, we developed a program to apply cloth simulation by optimizing the mesh of spatial mapping. As a result, it was possible to achieve a much more natural cloth rendering with spatial mapping mesh optimization comparing to the initial one. In the case of not optimized high resolution mesh output, the movement or collision of the cloth was relatively natural, but the result was a frame rate that could not be called real-time, and in the case when 60 percent of the mesh was optimized the frame rate maintained from 45 to 60 fps. With the current augmented reality hardware technology, it is reasonable to optimize the mesh and run the simulation to ensure real-time frame rate reliability. We are going to implement and study the deformable object's simulation using AR spatial mapping in various fields by using corresponding algorithm.

References

[1] Fujisawa, F.; Kato, H. Interactive Fluid Simulation Using Augmented Reality Interface; Springer: Berlin/Heidelberg, Germany, 2009.

[2] Jiménez, Pablo, Federico Thomas, and Carme Torras. "3D collision detection: a survey." Computers & Graphics 25.2 (2001): 269-285, 2001.

[3] Kim, M., Sung, N. J., Kim, S. J., Choi, Y. J., & Hong, M. Parallel cloth simulation with effective collision detection for interactive AR application. Multimedia Tools and Applications, 78(4), 4851-4868, 2019.

Suggestions of plans for revitalizing online-based education and the 4th industrial revolution technology development in the Cameroon environment

Sangwoo Hahm[1], Mbong Kome Ngole[1*], Isaac umalaka[1], Eta Ngole[2]
Department of Business Administration, Semyung University, Chungbuk, South Korea
Ministry of public health, Yaounde, Cameroon
[e-mail: thehahm@semyung.ac.kr, komembongn@gmail.com, i5446@yahoo.com, mbongeta@yahoo.fr]
*Corresponding author : Mbong Kome Ngole

Abstract

Developing countries, including Cameroon, are facing the challenge of using new technologies such as IT, 4th industrial revolution technologies (AI, Big data, and Cloud computing). These technologies will provide these countries with opportunities to create and develop new values. However, in reality, the educational system or infrastructure to learn and utilize these technologies from Cameroon and other countries is remarkably insufficient. This study, targeting Cameroon, seeks ways to revitalize online-based education and the 4th industrial revolution technology development.
This study proposes ways to improve revitalizing online-based education and 4th industrial revolution technology development through treating returning from their home country, investing use internet, building cooperative relationships with other countries in the context of Cameroon. Through these efforts, Cameroon will be able to achieve long-term development and creating new value opportunities to achieve national development.

Keywords*:* online-based education, 4th industrial revolution, technology development

1. Introduction

Developing countries, including Cameroon, are facing the challenge of using new technologies such as IT, 4th industrial revolution technologies (AI, Big data, and Cloud computing) [1]. These technologies will provide these countries with opportunities to create and develop new values. However, in reality, the educational system or infrastructure to learn and utilize these technologies from Cameroon and other countries is remarkably insufficient. This study, targeting Cameroon, seeks ways to revitalize online-based education and the 4th industrial revolution technology development.

2. Literature Review

First, in the case of Cameroon, there are insufficient schools or organizations that can educate these skills. Some students have to learn new skills by studying abroad. However, there is a limit to the success of all of these students, returning to their home country, using their skills and providing education to students in their home country. Therefore, there will be a need at

the national level to prepare necessary policies for the return of these international students, support them, and motivate them. The country should establish a student-friendly system for a study abroad life, and provide economic and emotional support to talented people. The government should recognize that returning to Cameroon is an opportunity to secure new resources and create new values [2].

Second, in Cameroon, except for Yaounde, the capital city, Internet penetration is low nationwide. The population of Cameroon is about 26,545,863 people, and the area of the country is 47544,000 ha. About 1,344,617 people (around 50% of the population), live in Yaounde, but the area is only 180 km². Therefore, the infrastructure concentrated in the capital will need to be developed by benchmarking from small countries such as Hong Kong and Singapore. Later, the expansion from Yaounde to other regions was concentrated in Seoul, the capital city, and Korea, which achieved development of other provinces together, or developed countries with a large area such as the United States [3]. Cameroon needs not only focus on short-term performance, but should invest in basic infrastructure such as Internet usage environment from a long-term perspective. Finally, Cameroon needs to revitalize its exchange program in cooperative relations with other countries. Some developed countries have friendly relations with African countries and support them. Cameroon is a country with various potentials including human resources. Hence, by establishing cooperative practices with other countries and learning various skills from them, Cameroon's capabilities could also be improved. To this end, Cameroon should maintain a more cooperative attitude and endeavor to operate as a trusted country.

3. Conclusions

This study proposes ways to improve revitalizing online-based education and 4th industrial revolution technology development through treating returning from their home country, investing use internet, building cooperative relationships with other countries in the context of Cameroon. Through these efforts, Cameroon will be able to achieve long-term development and creating new value opportunities to achieve national development.

References

[1] Mihyun Chung and Jaehyun Kim, “The internet information and technology research directions based on the fourth industrial revolution,” *KSII Transactions on Internet and Information Systems*, vol. 10, no. 3, pp. 1311-1320, March, 2016.

[2] Sangwoo Hahm, “Attitudes and performance of workers preparing for the fourth industrial revolution,” *KSII Transactions on Internet and Information Systems*, vol. 12, no. 8, pp. 4038-4056, August, 2018.

[3] Anand, Gurumurthy, and Rambabu Kodali. "Benchmarking the benchmarking models." *Benchmarking: An international journal,* 2008.

A Study on Hash Chain-based Authentication to Improve the Security of the VMF Military Standard

Dohoon Kim,[1] Sang Seo,[1] Heesang Kim,[1] and Won-Gi Lim[2*]
[1] Department of Computer Engineering, Kyonggi University, Suwon, South Korea
[2] Agency for Defense Development, Seoul, South Korea
[e-mail: karmy01@kgu.ac.kr, tjtkd8271@kgu.ac.kr, victoriousian@kgu.ac.kr, wklim@add.re.kr]
*Corresponding author: Won-Gi Lim

Abstract

Authentication algorithms in the form of an one-way cryptographic scheme, such as secure hash algori thm-1 (SHA-1) and the digital signature algorithm (DSA) specified in the current Variable Message Fo rmat (VMF) military standard, have numerous reliability-related limitations in a Tactical Data Link (T DL), which requires maximum security, communication integrity and low network overhead. In additi on, DSA-related RSA public key encryption, which is mainly used for VMF-based military communic ation authentication, also has the problem that it causes a slow processing speed and a large number of p rotocol header bits in performing rapid wired/wireless network communication within a dynamic battl efield environment with limited resources. Accordingly, in this study, a tactical wireless mobile ad-hoc network (MANET) topology, similar to that of the existing combat radio network (CRN), is considered, and a lightweight multi-factor hash chain for network overhead reduction and military terminal authent ication is proposed. The proposed method could enhance the confidentiality of message exchanges and reduce unnecessary network transactions and transmission bits for authentication flows between milita ry terminals owned by a squad, while ensuring robustness in a limited battlefield network, which is a m ajor issue in CRN. Based on these results, we intend in the future to increase the reliability between wi reless terminals in the KVMF CRN and apply it in-depth as lightweight authentication with exception h andling in Korea Army corps network scenarios.

***Keywords*:** Military; Variable Message Format (VMF); Hash Chain; Lightweight Secure Hash (LSH)

1. Introduction

Variable Message Format (VMF)[1] is a military digital information exchange standard established by the Joint Interoperability program under the United States Joint Chiefs of Staff. It provides common interoperability standards, including command data element and protocol standards, and it enables real-time transmission of digital tactical information in surveillance systems, C2 systems, and striking systems in a limited resource-based network-centric operational environment (NCOE).

However, since the tactical communication within the VMF-based CRN uses SHA-1-based DSA as the main authentication process, it can cause potential limitations in a low-speed and low-bandwidth dynamic network due to major problems such as practically proven hash-collision vulnerability, network overhead-based protocol header bits, and low transmission integrity. This will soon emerge as a negative security issue within the CRN in a TDL,

The authors gratefully acknowledge the financial support provided by the Defense Acquisition Program Administration and the Agency for Defense Development under the contract UD200023ED.

which requires maximum security, transmission integrity, and low network transmission overhead. In addition, such vulnerabilities will also become clear limitations when high-level adaptive digital covert activities and deception tactics are successfully implemented in hierarchical corps network-based network-centric warfare (NCW).

Furthermore, RSA encryption, which is mainly used for VMF authentication, will also have a separate problem that causes a slow processing speed, a large number of bits in protocol header and computational overhead in performing tactical information transmission in rapidly changing battlefields with limited resources. In other words, the existing VMF authentication must be improved to support the reliable transmission and reception of authentication information in limited CRN.

Accordingly, to solve these problems, a lightweight hash chain and related tactical scenario for network overhead reduction and military wireless terminal authentication-based MANET is proposed in this study.

2. Proposed Authentication Models

2.1 Design Principle

The proposed scheme was improved based on the multi-factor T-OTP hash chain like T/key[2] based S/key. The scheme is configured with an enhanced hash function, lightweight secure hash (LSH)[3], and Korean cryptographic hash standard to support issues such as hash-collision vulnerability, network overhead header bit, and low transmission integrity.

2.2 Lightweight Multi-Factor Hash Chain

The proposed processes consist of three phases: initialization and registration, authentication and verification, and re-authentication and revocation. The overall scheme for VMF is as shown **Fig. 1**.

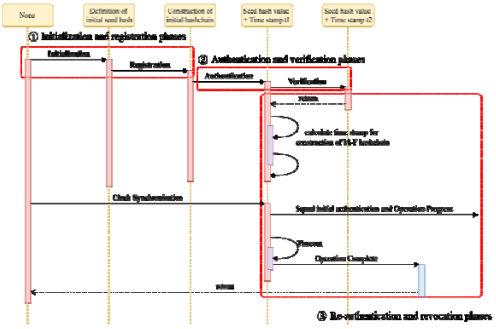


Fig. 1. Proposed Authentication Scheme in VMF

① Initialization and registration: OTP seed generation in an LSH-based hash chain between the administrator (military operation authority or squad head) and the user (connected device or squad member)

② Authentication and verification: fast authentication with timestamp-based clock synchronization and limitation of the lifetime of arbitrary hash values that have not been used for a long time

③ Re-authentication and revocation: tactical scenario-based regularization and exception handling in a CRN in a dynamic, changing battlefield network

After the initial authentication, the status of the current authentication progress is reflected in a real- time value such as serialization of a QR code, and the participating tactical equipment nodes are updated based on Korean Variable Message Format (KVMF) network environments. Subsequently, if an internal/external malfunctioning node or hostile node is detected and re-authentication is required, an additional authentication scheme can easily be executed at any time through clock synchronization of the hash functions in the LSH-based multi-factor hash chain of the joint operation group.

3. Experiments

The experiments focused on comparative analysis combined with compound military metrics and tactical operation scenarios. An overview of the NS-3[4] and MATLAB-based testbed with limited bandwidth in MANET is as shown in **Fig. 2**. The overall performance analysis, related results, and value types are as shown in **Fig. 3**.

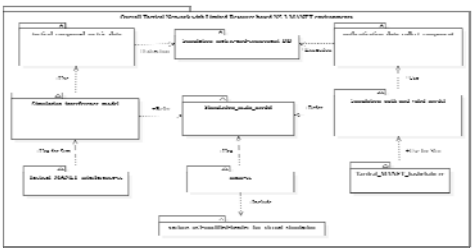


Fig. 2. Process flow in testbed with NS-3 environment

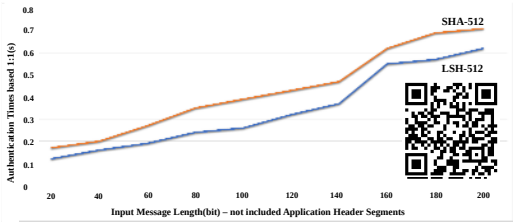


Fig. 3. Comparative authentication time results in limited bandwidth, lower 48Kbit/s, based MANET

4. Conclusions

In this study, a lightweight authentication-based LSH-hash chain applicable to a VMF-based CRN was proposed. This model could enhance the integrity of tactical message exchanges and reduce unnecessary network transactions and transmission bits for authentication flow in an arbitrary CRN, while ensuring robustness in a low-speed and limited-bandwidth network. In the future, we intend to increase the reliability between wireless devices in the KVMF and apply it in-depth as lightweight authentication in Korea Army network scenarios.

References

[1] MIL-STD-2045/47001D (w/CHANGE1), Department of Defense Interface Standard: Connectionless Data Transfer Application Layer Standard, June 2008,

[2] D. Kogan, N. Manohar, and D. Boneh. "T/key: second-factor authentication from secure hash chains." *in proc. of the ACM SIGSAC on Computer and Communications Security*, 983-999, 2017.

[3] K. Dong-Chan, et al. "LSH: A new fast secure hash function family." *International Conference on Information Security and Cryptology*, 286-313, 2014.

[4] NS-3, Available online: https://www.nsnam.org/

A Practical Data Preparation Approach for Machine Learning over Heterogeneous Data Sources

Taewhi Lee, Jang-Ho Choi, Miyoung Jang, Jongho Won, and Jiyong Kim
Smart Data Research Section, AI Research Lab.
Electronics and Telecommunications Research Institute (ETRI)
Daejeon, Republic of Korea
[e-mail: {taewhi, janghochoi, myjang, jhwon, kjy}@etri.re.kr]
*Corresponding author: Taewhi Lee

Abstract

Machine learning applications often involve various kinds of data from heterogeneous data sources. It is tedious, time-consuming, and error-prone to integrate such heterogeneous data, so we need an easy-to-use solution. This paper presents a practical approach to prepare data for machine learning that uses heterogeneous data from multiple sources. This approach can support federated queries over heterogeneous data sources and multi-dimensional array operations for data transformation. Such queries can be defined as views for user convenience. Users can therefore query the integrated data with some arbitrary conditions in a convenient manner.

***Keywords*:** machine learning data preparation, heterogeneous data integration

1. Introduction

When analyzing data using machine learning, it is often necessary to analyze multiple datasets together. For example, several datasets are involved in the machine learning process to build prediction models for green algal blooms, such as water quality data, meteorological data, hydrological data, and flow data [1]. As data need to be analyzed from various perspectives, it may be required to prepare a large number of datasets which are filtered with different conditions. These data preparation tasks are important but very cumbersome and time-consuming to implement.

Our goal is to easily get various kinds of datasets in the form of tensors (multi-dimensional arrays), which are the basic data structures in machine learning. There have been attempts to make data preparation convenient for machine learning. Kim and Candan proposed TensorDB [2], which support relational and tensor algebraic operations on chunk-based array data stores by leveraging array operations. Leclercq and Savonnet proposed a Tensor Data Model [3], which is logically independent from a particular data stores. As the name suggests, this model is based on tensors and defines a set of tensor operators. They presented this model in a polystore system architecture.

Inspired by these studies, we present a practical data preparation approach for machine learning in this paper. Our approach can help users to retrieve their datasets in the form of tensors from different data sources via simple queries.

This work was supported by the Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No. 2018-0-00219, Space-time complex artificial intelligence blue-green algae prediction technology based on direct-readable water quality complex sensor and hyperspectral image).

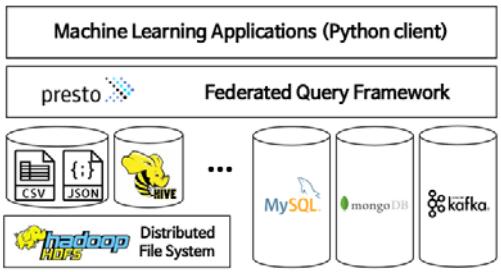


Fig. 1. System architecture

2. System Architecture

Our system requirements can be summarized as follows: (1) Support of tensor data manipulation. (2) Convenient data retrieval for application users over heterogeneous data sources.

Fig. 1 shows our system architecture for tensor data management. We employ Presto[1] as our federated query framework, which is an open source distributed query engine that currently supports over 20 data sources. It also supports array functions that are useful for data manipulation.

Consider an example scenario for algal bloom prediction on our study area (Daecheongho lake). Suppose that two kinds of data, such as weather data and water quality data, are used to train a CNN model and they are stored in different data stores as in **Fig. 2**. Such data can be transformed into the input tensor form of the CNN model via federated queries using array functions. Administrators can create views as follows so that application users can use them in a convenient manner, as shown in **Fig. 3**.

```
CREATE VIEW hive.abap.dch_tensor AS
WITH  t AS (
  SELECT ngrams(array_agg(q.day), 3) day,
          ngrams(array_agg(water_temp), 3) water_temp,
          ngrams(array_agg(insolation), 3) insolation
  FROM mysql.abap.dch_water_quality q,
         hive.abap.dch_weather w
  WHERE q.day = w.day )
SELECT dch.*
FROM t, UNNEST(
   t.day, t.water_temp, t.ph, t.ec, t.do, t.chl_a, t.toc,
   t.tn, t.tp, t.avg_temp, t.min_temp, t.max_temp,
   t.rainfall, t.insolation)
  dch(day, water_temp, ph, ec, do, chl_a, toc, tn, tp,
   avg_temp, min_temp, max_temp, rainfall, insolation);
```

Our apporach is easy to use for application users, and practical to implement for data administators.

[1] https://prestodb.io/

```
presto> select * from hive.abap.dch_weather;
   day     | avg_temp | min_temp | max_temp | rainfall | insolation
----------+----------+----------+----------+----------+------------
 7/1/12    |     22.9 |     20.1 |     27.7 |      0.0 |      15.25
 7/2/12    |     23.6 |     19.5 |     28.2 |      0.0 |       22.4
 7/3/12    |     24.1 |     19.1 |     29.0 |      0.0 |      15.45
```

(a) Weather data (CSV files or Hive external tables)

```
presto> select * from mysql.abap.dch_water_quality;
   day     | water_temp |  ph  |  ec   |  do  | chl_a | toc |  tn   |  tp
----------+------------+------+-------+------+-------+-----+-------+-------
 1/1/13    |        6.7 |  7.5 | 120.0 | 10.8 |   3.7 | 2.1 |   1.4 |  0.01
 1/1/14    |        7.7 |  7.7 | 152.0 | 11.1 |   2.7 | 2.0 | 1.877 | 0.012
 1/1/15    |        7.5 |  7.6 | 141.0 | 10.8 |   3.5 | 2.2 | 1.411 | 0.007
```

(b) Water quality data (MySQL tables)

Fig. 2. Example scenario

```
In [83]: conn = prestodb.dbapi.connect(host='localhost',port=8080,
                    user='etri',catalog='hive',schema='abap')
         cur = conn.cursor()
         cur.execute('select * from dch_tensor')
         rows = cur.fetchall()

In [84]: presto_data = pd.DataFrame(rows)
```

Fig. 3. Usage in machine learning applications

Users can retrieve their data with arbitrary conditions in the tensor form. Also, each dataset can be managed by different entities. Since views are not physical objects, preparation time can be reduced by materializing frequently used data.

3. Conclusion

In this paper, we present a practical data preparation approach for machine learning over heterogeneous data sources. The proposed system can provide users more convenient data transformation for machine learning.

References

[1] T. Lee, J.-H. Choi, M. Jang, J. Won, and J. Kim, "Enhancing Prediction of Chlorophyll-a Concentration with Feature Extraction using Higher-Order Partial Least Squares," in *Proc. of International Conference on ICT Convergence (ICTC)*, pp. 1666-1668, 2020.

[2] M. Kim and K. S. Candan, "TensorDB: In-Database Tensor Manipulation with Tensor-Relational Query Plans," in *Proc. of ACM International Conference on Information and Knowledge Management (CIKM)*, pp. 2039-2041, 2014.

[3] E. Leclercq and M. Savonnet, "TDM: A Tensor Data Model for Logical Data Independence in Polystore Systems," Heterogeneous Data Management, Polystores, and Analytics for Healthcare - VLDB 2018 Workshops, Poly/DMAH, pp. 39-56, 2018.

KSII The 12th International Conference on Internet (ICONI) 2020.

A study of building techniques for easy and fast implement an application in various mobile environments

Kunjung Sim and LIN BIN and YoungHwan Lim
Department of Media, Soongsil University
Seoul, South Korea
[e-mail: inyourhands@nate.com, linbin98@gamil.com, yhlim@ssu.ac.kr]
*Corresponding author: YoungHwan Lim

Abstract

In the rapid growth and mature environment of mobile applications, the demand for the development and distribution of mobile applications is becoming increasingly diverse. However, the development period of mobile applications is still not as short as expected, and there are considerable difficulties in applying them to actual mobile devices even after the planning and design are completed.
To solve these problems, this paper proposes the following four building techniques. First, users should be able to configure their mobile applications themselves. Second, the design of mobile applications and UI-UX should be able to be quickly configured in the form desired by users. Third, mobile applications should be able to be deployed quickly so that users can deploy them at any time they want. Fourth, it should be able to deploy to a variety of mobile OSs and different mobile devices. Based on the above four building techniques, the purpose of this study is to enable users to quickly and easily create mobile applications so that anyone can deploy them in a timely manner at a relatively low cost.

***Keywords*:** Mobile App, Smartphone App, Building Techniques, Hybrid Mobile Application, Cross-Platform Mobile Application

1. Introduction

Despite the steady growth of the mobile app market, it is still not easy to develop mobile apps. This is because mobile manufacturers are manufacturing devices based on different platforms and OS.

There have been various efforts to develop Cross-Platform Mobile Application[1] to solve these problems. Based on the Hybrid Mobile Application[2] development method, which has received much attention until recently, this paper researches and proposes a build-up technique that allows ordinary users to plan, develop and distribute mobile apps easily and quickly.

2. Building Techniques

2.1 Real-time Prototyping

Real-time Prototyping is based on the development method of Hybrid Mobile Application, which is one of various mobile application development methods. Hybrid Mobile Application development method is based on WebView which is provided by almost all mobile platform, and UI/UX are configured using HTML5, CSS3, and Javascript, and application is developed by linking core functions provided by Native through Plug-in[4].

Real-time Prototyping can reduce the process of building and porting each time by using the existing method, because the UI/UX can be checked in real time through the web browser

which is an advantage of the method of developing the Hybrid Mobile Application[5].

2.2 Fast Customizing

Fast Customizing requires existing page templates and function module templates to configure resource assets separately, allowing users to break away from the same shape and form of configuration by modifying the resources of the desired page or function module.

Changes to these resource assets are immediately reflected and can be viewed directly from a web browser via Real-time Prototyping emulator.

2.3 In-time Publishing

The most complex and difficult part of how to develop existing mobile apps is building source code for various mobile platforms. Mobile device manufacturers are each dependent on their own language and devices to develop apps, making it difficult for end users to solve them.

In-time publishing increases deployment immediateness by enabling the deployed back-end server to build deployment files for the desired mobile platform at any point in time.

2.4 Multi-OS, Multi-Platform Support

To solve Multi-OS, Multi-Platform problems in the building technique, build back-end servers through the Apache Cordova framework, one of the Hybrid Mobile Application Development frameworks. The Apache Cordova framework now supports Android, iOS, Windows Phone 8.1, and 10 major smartphone platforms, as well as Linux, Windows, and OS X, which are PC platforms, so you can build source code without having to configure all platform build environments through the back-end servers that have the Apache Cordova framework.

3. System Consists

Fig. 1. System Consists

Building techniques system consists largely of prototyping web services for real-time planning and real-time design changes, and backend servers for building source code and download files for each store distribution(**Fig. 1**).

4. Conclusions

Building techniques studied in this paper is based on the development method of Hybrid Mobile Application, which is one of mobile application development methods. It is based on Real-time Prototyping technique that enables to check the appearance of the developed application without actually porting to the actual device, Fast Customizing techniques that allow to change quickly, In-time Publishing techniques that allow to build in the right time to download into the App Store and download distribution files, and Multi-OS, Multi-Platform techniques, making it possible for ordinary users, not professional mobile app developers, to quickly and easily create mobile apps.

References

[1] Popa, Marius, "Considerations Regarding the Cross-Platform Mobile Application Development Process," *Academy of Economic Studies. Economy Informatics; Bucharest13.1 (2013): 40-52.*

[2] Isabelle Dalmasso, Soumya Kanti Datta, Christian Bonnet, Navid Nikaein, "Survey, Comparison and Evaluation of Cross Platform Mobile Application Development Tools," *Wireless Communications and Mobile Computing Conference (IWCMC), 2013 9th International.*

[3] Christos Bouras, Andreas Papazois, Nikolaos Stasinos, "A Framework for Cross-platform Mobile Web Applications Using HTML5," *International Conference on Future Internet of Things and Cloud, 2014.*

[4] Stefan Bosnic, Ištvan Papp, Sebastian Novak, "The development of hybrid mobile applications with Apache Cordova," *Telecommunications Forum (TELFOR), 2016 24th.*

[5] Latif, Mounaim, Younes Lakhrissi, and Najia Es-Sbai. "Cross platform approach for mobile application development: A survey." *2016 International Conference on Information Technology for Organizations Development (IT4OD)*. IEEE,.

Region-based video software game sharing application

Yoseb Lee and YoungHwan Lim and Kunjung Sim and LIN BIN
Department of Media, Soongsil University
Seoul, South Korea
[e-mail: hehe82@nate.com, yhlim@ssu.ac.kr, inyourhands@nate.com, linbin98@gmail.com]
*Corresponding author: YoungHwan Lim

Abstract

Expose the way you share and share your video software games based on your location base, selectively reserve and rent video software games based on your area, check available video software games in your area, rent them or sell them, and automatically present the price of video software games using data in the existing market price range, and notify the amount of video software games to allow users to use.

***Keywords*:** Video game, Regional based, share, application

1. Introduction

Recently, the market of console game consoles is expanding. Console game consoles such as PS4, XBOX One, and Nintendo Switch were released, and the market for console games began to grow. As the market for console games grows[1], digital download games have emerged, making it easy to download and play games, but games with more than 10 million console game CDs sold[2] in 2019 are also being released. If you look at the ending of the console game, the value of the game decreases, so it is difficult to dispose of the console game CD.

In order to solve this problem, we are developing a regional-based console game sales service.

To provide regional-based console game sales service, the following functions must be provided.

1) Tradeable areas within 10km of your area are marked.

2) Adjust the price of products sold through the game auction system.

3) We provide a reservation system to reserve transactions.

2. Features provided

2.1 Transaction area search function

In order to buy and sell console game CDs, you need to set the region where you are selling.

It provides information on console game CDs sold within 10KM[3] of their region through GPS. Based on your location, the product of the console game CD is exposed within 10KM, and the lowest price and seller information of the set amount according to the proposed amount of the game auction system are exposed.

Fig. 1. Region-based sales list

2.2 Console game pricing auction system

The selling price of console game CDs is important. In order to set the price of console game CDs, it is important how many console game CDs are on sale. When the user registers the information of the console game CD in order to sell the console game CD, the system will propose the final price of the currently registered console game CD, and the user selects the desired amount from the proposed amount and the amount of the console game CD You choose. Through this system, a buyer can purchase a console game CD at a differentiated price according to the supply amount of the console game CD.

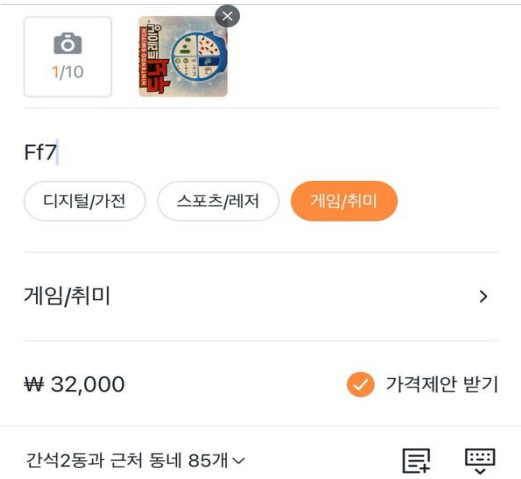


Fig. 2. Region-based console game sales information registration

2.3 Transaction reservation system

To purchase a console game CD, you need to make a deal with the seller. In order to proceed, you need to designate a trading place and trading time. The principle is that the buyer moves to the seller's place. To do this, the trading hours between the seller and the buyer must be set. When the buyer wishes to make a transaction, the seller suggests the time available for sale through the sales system. The buyer sees the suggested time and makes changes or approvals, and the buyer goes to the seller's desired location and purchases a console game CD.

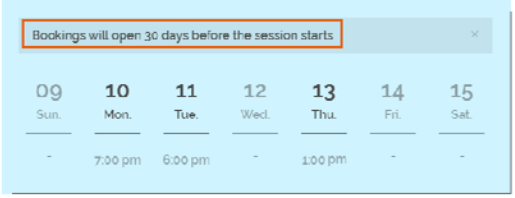


Fig. 3. Designation of transaction date for regional console game CD transaction

3. Conclusions

Through the region-based video software game sharing application, buyers and sellers will be able to trade console game CDs to be sold and console game CDs to purchase at an appropriate price with a person at their location. The seller can set the desired price for the console game CD and can easily grasp the flow of how the current price is going. Buyers will have more opportunities to purchase console game CDs they want at low prices.

References

[1] BLOG [Internet]. Available: https://blog.naver.com/kyjoo031216/222131748258

[2] VGChartz [Internet]. Available: https://www.vgchartz.com/weekly/43464/Global/

[3] Region-Query [Internet]. Available: https://medium.com/sunhyoups-story/gps-%EC%A7%80%EC%97%AD-%EC%B0%BE%EA%B8%B0-%EA%B4%80%EB%A0%A8-%EC%BF%BC%EB%A6%AC-789f6200b96c

Runtime Memory Usage Optimization For 64-bit OS

Jinman Kwon, Chungki Woo and Jaeook Kwon
Samsung Electronics Co., Ltd.
Suwon, South Korea
[e-mail: jm0508.kwon@samsung.com, chungki0201.woo@samsung.com, jook.kwon@samsung.com

Abstract

Most operating systems widely choose 64-bit architecture due to its larger address space & general expectation that its performance would be faster than that of 32-bit architecture. However, 64-bit architecture tends to increase memory usage, thus degrading its performance.
Here, we aim to find the reasons of the increase in runtime memory usage when the 64-bit kernel is used instead of 32-bit kernel. And we suggest strategies to use physical memory efficiently by optimizing the kernel structure and decreasing slab allocation size per structure. We demonstrate how to minimize the amount of increase in runtime memory, which is caused by changing the architecture from 32-bit to 64-bit, via utilizing several "Slab struct optimization methods". The result shows about 28.9% memory reduction.

Keywords: 64-bit, Memory, Slab allocator

1. Introduction

The 64-bit Linux architecture has been adopted and by many operating systems such as Google Android and Web OS. The 64-bit OS can scale memory address and perform faster than the 32-bit OS. However, 64-bit architecture uses more memory compared to 32-bit architecture, thus its performance can be deteriore free memory space may become insufficient in products equipped with the same DDR RAM when the 64-bit OS installed.

Issues related to the increase of memory usage when a 64-bit OS is chosen for Linux development have been often neglected. The reason for the increase of memory usage is that the Linux OS open source adheres to the versatility to run on a variety of devices including server products. This memory increase may be a tiny issue in server products because huge size of RAM is installed in them, but since embedded systems have limited RAM, a small increase in memory usage can have a significant impact on the overall system operation.

In the 64-bit OS, memory usage can increase in both user and kernel space, and there are several reasons for this phenonmenon. We intensively studied the kernel, especially on the increase in the use of runtime memory of the kernel instead of kernel reserved memory such as memmap. It is because we observed that about 79% of the total kernel memory growth increases in runtime memory.

Section 2 shows the increase of runtime memory at 64-bit kernel compared to 32-bit kernel, which explains why the runtime memory was studied. In section 3, we suggest how to save the runtime memory.

2. Runtime memory usage increasing in 64-bit OS

68% of the total increase of kernel runtime memory was caused by the slab memory usage. As the increase in slab usage is likely to degrade performance due to insufficient memory, the use of slab memory needs to be improved.

Many operating systems using the Linux kernel mainly choose the slab allocator as a runtime memory allocation method. Slab allocation is a memory management mechanism intended for the efficient memory allocation of objects. [1] We had found that the 64-bit kernel uses 49.9% more slab memory than the 32-bit kernel when measured on Tizen OS 5.5 using the same source codes. It is because 64-bit architecture has a feature that increases slab usage.

The feature of 64-bit kernel changes the size of the long type and pointer variable. The long type's variable or the pointer variable use 4 bytes in the 32-bit architecture, whereas 8 bytes are used in the 64-bit architecture. Therefore, if a structure contains a variable of type long or a pointer variable, it requires a larger slab memory allocation in a 64-bit kernel environment.

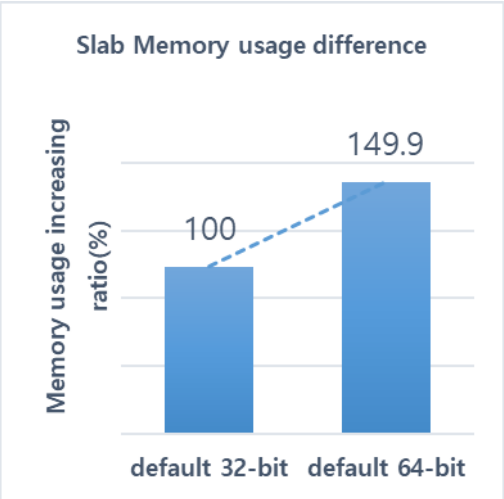


Fig. 1. Slab usage between 32-bit and 64-bit

3. Runtime memory optimization methods

Based on the cause that increases memory usage mentioned in Section 2, we suggest five solutions below for application.

1) Change variable type from long to int
2) Do union pointer variable and another
3) Utilize bit field
4) Change variable positions for byte padding
5) Avoid additional cache line alignment

1) ~ 3) are to reduce the size of members in the structure considering the characteristics of embedded products. 4) ~ 5) are to optimize size for structures in whiche alignment size has changed as the size of member variable increases.

3.1 Variable type change from long to int

In 64-bit architecture, compared to 32-bit, long type variables occupy 4 bytes more, and the amount of slab memory usage increases. Among the variables in these structures, if you find variables that use only the size within 32-bit and change to int type instead of long type, you can change it to have the same slab memory usage as in 32-bit architecture.

For example, in the case of the address_space structure, it is possible to save 16 bytes by changing the variable to an int type after verifying that there is no value exceeding 32-bit in the investigation of 'nrpages' and 'nrshadows' in the embedded system with RAM below 4Gbytes, it was possible to save 16 bytes by changing it to an int type.

Table. 1. Suggested change of variables type from long to int

As-is	To-be
struct address_space { … **unsigned long** nrpages; **unsigned long** nrshadows; …	struct address_space { … **unsigned int nrpages;** **unsigned int nrshadows**; …

3.2 Merging a pointer variable

When a pointer variable stores an address, it should allocate 8 bytes in a 64-bit architecture. However, if The pointer's address in 64-bit starts with fixed area like 0xfffff0000000000, another variable can use some bit areas of the pointer variables.

In the case of the pointer variable 'owner' in the

'rwlock' structure, it is sufficient to use only the address in the 56-bit range in the actual embedded system, so the code was modified so that the remaining 8 bits are used by other variables.

Table. 2. Merging pointer variable and another

As-is	To-be
typedef struct { … #ifdef CONFIG_DEBUG_SPINLOCK unsigned int magic, **owner_cpu;** **void *owner;** #endif … } rwlock_t;	typedef struct { ... #ifdef CONFIG_DEBUG_SPINLOCK unsigned int magic; **unsigned long long owner_cpu:8;** **unsigned long long owner:56;** #endif … } rwlock_t;

3.3 Utilizing bit field

Configuring structure members by using bit fields in consideration of alignment also helps to reduce the size of the structure.

Since the 'signum' and 'pid_type' members do not exceed the value of 2 bytes- size in the embedded system environment with little sized RAM in small size, the size of the structure can be decreased by changing the data type and position of two members according to the alignment.

Since the size of 'rwlock_t' has reduced to 4 bytes in 3.4, the alignment of the 'signum' and 'pid_type' members was aligned.

Table. 3. Utilizing bit field

As-is	To-be

struct fown_struct {
rwlock_t lock;
struct pid *pid;
enum pid_type pid_type;
kuid_t uid, euid;
int signum;
};
	struct fown_struct {
	rwlock_t lock;
	signed short signum;
	enum pid_type pid_type:16;
	struct pid *pid;
	kuid_t uid, euid;
	};

3.4 Changing positions for byte padding

Structures that consider only 32-bit architecture may not consider byte padding[3] compiled with 64-bit architecture. This mistake results in an increase in the actual memory size of the structure as the alignment of the structure changes. These structures can rearrange their members in consideration of alignment in a 64-bit architecture, thereby eliminating or minimizing additional byte padding compared to 32-bit. Since such byte padding issues can change according to the enabling/disabling of the kernel config, it is necessary for developers to comprehensively consider whether configs related to each structure are enabled or not.

In the case of the "struct mutex" structure, it is possible to prevent additional memory allocation due to byte padding by reducing the *owner to 4 bytes and changing the positions in the wait_list and wait_lock structures.

Table. 4. Change positions for byte padding

As-is	To-be
struct mutex { /* 1: unlocked, 0: locked, negative: locked, possible waiters */ atomic_t count; **spinlock_t wait_lock;** **struct list_head wait_list;** …	struct mutex { **struct list_head wait_list;** **spinlock_t wait_lock;** /* 1: unlocked, 0: locked, negative: locked, possible waiters */ atomic_t count; …

3.5 Avoiding needless cache line alignment

Linux kernel operates fast by aligning cache line[3]. As a result, each structure occupies a larger memory space than the size specified in the code. Since the size of the padding is set based on the alignment, if the alignment is exceeded even by a little, the structure size increases significantly. Such cache line alignment is utilized when only pure source codes are considered excluding the alignment by the compiler, or when only the characteristics of the existing 32-bit architecture are focused on. Therefore, it is necessary for a developer to comprehensively judge and code the above various matters when defining a variable for a structure.

In the 'inode_info' structure, 64 bytes can be saved in total in the structure, although it appears only 8 bytes can be saved in code when the position of the 'open_count' variable is adjusted. , it seems that we can just save 8 bytes in code, but the actual total size saving of the structure is 64bytes.

Table. 5. Moving variable to avoid bytes padding

As-is	To-be

struct inode_info {
struct inode vfs_inode;
…
struct list_head orphan_list;
__u64 next_orphan_id;
atomic_t open_count;
#ifdef
…
};
	struct inode_info {
	struct inode vfs_inode;
	__u8 record_type;
	atomic_t open_count;
	unsigned long flags;
	…
	};

4. Measurement result

The actual implementation and evaluation in this paper was performed using Samsung Tizen OS.

Table. 5. Moving variable to avoid bytes padding

Embedded system test board	
CPU	Cortex-A73 (4 cores)
OS	Tizen 5.5
Kernel version	v.4.1

As a result, we successfully saved memory from 3 structures out of 33 kernel slab objects and were able to reduce the memory usage increase rate from 149.9% to 135.5% in the 64-bit kernel-based test board environment. This result means that we saved 28.9% of the increasing slab memory by changing to 64-bit architecture.

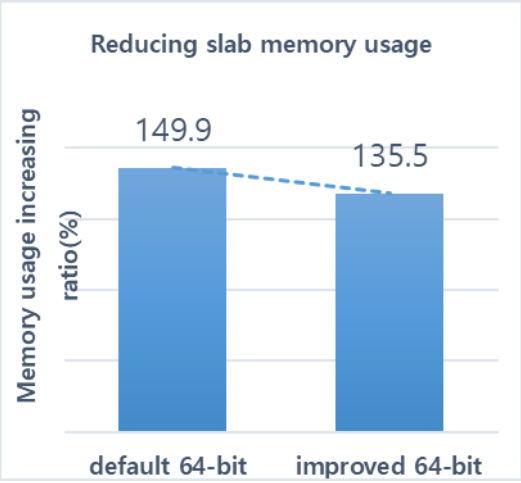


Fig. 2. Overall memory usage reduction effect after structure size improvement

5. Conclusions

In this study, we identified the causes of the increase in the slab memory capacity among the factors affecting the runtime memory in the 64-bit Linux OS architecture. We also demonstrated several ways to reduce memory usage without compromising performance.

By reducing the size of the kernel structure and minimizing padding, it was possible to show a smaller size of runtime memory usage compared to the default 64-bit Linux kernel. With the current development trend, it is expected that the runtime memory usage of the OS will continue to increase due to the installation of various new functions. Therefore, we plan to continue research on_other memory saving methods and applying our methods to another products including server products in addition to this study in the future.

References

[1] Jeff Bonwick,The Slab Allocator: An Object-Caching Kernel Memory Allocator ,1994.

[2] Wolfgang Mauerer, "*Professional Linux Kernel Architecture*", 2010:Wrox

[3] Daniel Pierre Bovet, Marco Cesati, "*Understanding the Linux Kernel*", PP.248, 2005:OReilly

Optimization Of Interprocess Communication For Embedded OS

Yonghyeok Lee, Chungi Woo and Jaewook Kwon
Samsung Electronics Co., Ltd., Korea
[e-mail: yhji.lee@samsung.com, chungki0201.woo@samsung.com, jook.kwon@samsung.com]

Abstract

Most embedded OS (Android, Tizen, etc.) supports multi-process and provides interprocess communication (IPC) techniques for process communication. These IPC techniques can cause a problem of over-occupying system resources if they are used inappropriately for their purpose.

In this paper, we have investigated these IPC techniques and real usage in Tizen OS. As a result, we found the problem that is being misuse. Finally, we have improved the problem and solve the problem of excessive occupancy of system resources. All experiments and measurements were successful, we decrease 10% of threads in total system.

***Keywords*:** IPC (Interprocess Communication), Tizen, Android

1. Introduction

Unix-based operating system supports various types of interprocess communication. These IPC techniques are also used in various forms in the Embeded OS. Typically, Android's binder, ios' mach message, and Tizen's dbus are provided for the IPC.

These IPC techniques are used a lot in the system. Also because IPC use system resources, It can cause a problem of over-occupying system resources if they are used inappropriately for their purpose.

Therefore, it is very important to understand the architecture of IPC techniques and use them correctly.

In this paper, we have investigated these IPC techniques and their uses. In the Section 2, we have investigated the features of IPC techniques used by Embedded OS. After that, we checked the usage type of IPC technique of one of OSs. Finally, we found the problem of over-occupying system resources by using it in an inappropriate way. In Section 4, we improved this and measured the system resources before and after the improvement.

2. Research

First, we have investigated architecture of IPC on Android and Tizen OS among embedded OS.

Android supports binder framework for IPC. IBinder is base class for a remotable object, the core part of a remote procedure call mechanism. [1]. If there are two process and they trying to communication each other, receiver process should set address to receive request from sender. Also receiver has a thread pool for handling on requests. The Binder framework will spawn new thread from thread pool to handle all incoming requests. It means, receiver process can handle multiple requests at the same time, for which multiple threads run simultaneously [2]. Because of this architecture, there is a possibility that system resource consumption may increase.

In case of IPC in Tizen OS, there are similar parts to Android. GDBusConnection is base object for remote call like Android's IBinder. When communicate within two process, the

receiver process set the address to this connection to receive message. And a worker thread is also created to handle request. But unlike the IBinder, receiver have worker threads each address not for each request and sender has additional worker thread for asynchronous call. So we have to be careful for problem that too much worker threads are created for IPC in Tizen OS. In the next section investigated real usage of IPC on Tizen OS relate to this worker thread problem.

3. Usage of IPC in Tizen

The worker thread model of Tizen IPC depends on the connection type. GDBusConnection have two types of shared and private, which are classified according to their use. Shared type is the default connection of the process. Both the sender and receiver use this connection to communicate and a worker thread is created. Private type is used when receiver needs extra address. A new worker thread is created to handle request for new address.

However we found a problem with some sender use private connection for send message, while checking the usage of IPC in Tizen OS. In this case, this private connection doesn't have their own address for receive. Due to this problem, as shown in the graph below, a number of unnecessary worker threads were created. As a result, system resources are occupied excessively.

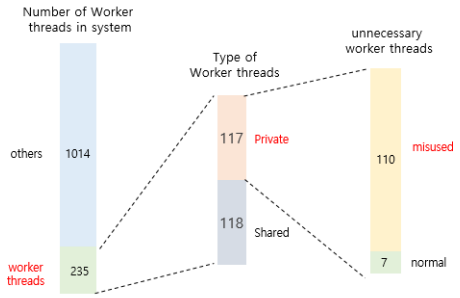


Fig. 1. Number of worker threads in Tizen (before)

Fig. 1 shows the number of worker threads in the whole system and type of worker threads. About 20% of threads are worker thread for IPC and half of them were made from private connections. We could even check that most private connections were made by senders.

Next section, we verify theses connections and modify if misused.

4. How to Find and Improve

To improve this problem, we have to checked whether the private connection uses a their own address. Otherwise, it was used incorrectly for the purpose, so we can change it to a shared connection. [3]

At first, we found a process with more than two worker threads in Tizen OS. (**Fig. 2**)

```
sh-3.2# ps -eL | grep 570
  570   570 ?      00:00:00 processA
  570   728 ?      00:00:00 gmain
  570   730 ?      00:00:00 gkdbus
  570   915 ?      00:00:00 gkdbus
  570   916 ?      00:00:00 gkdbus
  570   916 ?      00:00:00 gkdbus
```

Fig. 2. Processes with one or more workers

Next, to verify if worker thread have an address, we used tool named 'busctl' that used to introspect and monitor the DBus. [4]

```
sh-3.2# busctl   | grep 570
:1.67                570
:1.69                570
:1.75                570
org.tizen.processA   570    // have an address
```

Fig. 3. Check Connection's address Using busctl

Fig. 2 shows four worker threads created to handle request. But, **Fig. 3** shows that only one connection has an address. It means there are three unnecessary worker threads.

```
-address = g_dbus_address_get_for_bus_sync(...)
-con = g_dbus_connection_new_for_address_sync(...)
+con = g_bus_get_sync()
...
```

Fig. 4. Change Connection Type Private to Shared

So we could change three connections to shared connection like **Fig. 4**.

```
sh-3.2# ps -eL | grep 582
  582   582 ?      00:00:00 processA
  582   790 ?      00:00:00 gmain
  582   816 ?      00:00:00 gkdbus
```

Fig. 5. Result after Connection Type Change

As shown in **Fig. 5**, the worker thread has decreased after change of connections's type.

Next section, we applied it to the entire system and measured the results.

5. Experiment and Result

Table. 1. Test Environment

Embedded system test board	
CPU	Cortex-A73
OS	Tizen 6.0

After change the type of connection, we decrease about 100 connections and the result is as follows.

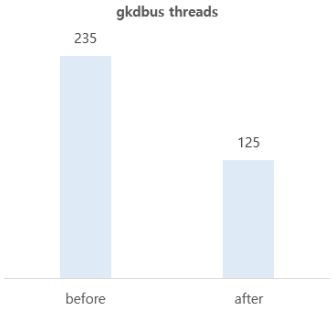


Fig. 6. Number of worker threads in Tizen

Fig. 6 shows the decrease of the number of worker threads. It was reduced by about 55%. Since thread uses system memory resources, we also checked how much memory usage improves by reducing the number of threads.

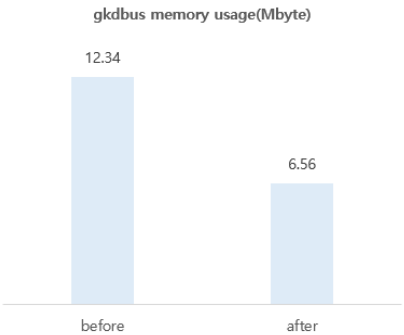


Fig. 7. Changes in IPC memory usage

From **Fig. 7**, we can see that the memory is reduced by about 6M after removing the worker thread.

6. Conclusion and Future work

As a result of analyzing and improving the IPC technique and architecture, system performance was optimized such as 10% reduction in threads and 6MB memory reduction. Most IPC techniques use separate worker threads, memory, etc. for performance and responsiveness. Android's binder and ios' mesh message has a similar in architecture. Therefore, if we understand about IPC techeniques and use it properly, we can greatly optimize system performance in other IPC techniques.

References

[1] T. Schreiber, J. Somorovsky, “Android Binder” , 2011.

[2] “Process and threads overview” in Android Development Documents, https://developer.android.com/guide/components/processes-and-threads

[3] “GIO Reference Manual”, gnome developer https://developer.gnome.org/gio/stable/GDBusConnection.html

[4] “busctl”, system , freedesktop.org, https://www.freedesktop.org/software/systemd/man/busctl.html

Blockchain-based AI Malware Detection Knowledge-Sharing and Private Data Market Platform

Uk Jo, Dohun Kim and Howon Kim
School of Computer Science and Engineering, Pusan National University
Pusan, South Korea
[e-mail: jouk@islab.re.kr, dohun@islab.re.kr, howonkim@pusan.ac.kr]
*Corresponding author: Uk Jo

Abstract

In recent, mobile use has been increased rapidly, and then the many platforms have been developed to be easily used in mobile. As well as people often use the mobile to do simple things such as games, also people deal with private data or company critical information such as banking service, business chat, and company business in mobile, Etc. The increase in mobile use has created an opportunity to move the target of malicious app developers to mobile, and the number of malware apps targeting mobile is increasing exponentially every year. Therefore, coping with malicious apps that are poured out every year is difficult. For these problems, the key point is collective intelligence through knowledge sharing, and at the same time, it is also necessary to grant the value of the own private data used for solving the problem. Therefore, this paper proposes the blockchain-based AI knowledge sharing and knowledge market platform for malicious app detection. In conclusion, each AI malware detection program has improved its performance, and the proposed platform architecture securely transacts private data.

***Keywords*:** Blockchain, Malware App, Malware Detection, Private Data Market, Sharing AI Knowledge

1. Introduction

The use of a mobile phone is increasing, and the number of malicious apps increases exponentially every year. Many people and organizations have developed malware app detection programs, but this program is limited to individuals' knowledge or relatively small organizations' knowledge in detecting the malicious app. If various developers share their knowledges, they will be able to deal with malicious apps that increase every year. But this leads to another problem where personal private data cannot be guaranteed. Therefore, this paper presents a platform that enables value transactions of private data of individuals and organizations based on blockchain, and that can solve the problem of malicious apps by sharing knowledge and receiving feedback among participants in the blockchain network. This platform has two purposes. First, it aims to solve the poor performance when individuals or relatively few organizations detect malicious apps by sharing knowledge with people participating in the platform. Second, participants have the monetized digital private data by using the blockchain

This work is financially supported by Korea Ministry of Land, Infrastructure and Transport(MOLIT) as 「Innovative Talent Education Program for Smart City」
This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2020-0-01797) supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation)

2. Blockchain-based AI Malware Detection Knowledge-Sharing and Private Data Market Platform

2.1 Architecture

Blockchain is a decentralized distributed ledger-based technology.[1] When blockchain-network has the transaction, it creates blocks containing transaction data that includes the hash value of the previous block to connect with the previous block. The smart contract(chaincode), when certain conditions are satisfied, makes nodes perform the prescribed policies automatically.[2,3] In particular, blockchain allows people to get the monetized digital data that can be traded with cryptocurrency.[4]

The proposed platform has two parts, off-chain that store various data and on-chain where the blockchain network is built. The off-chain consists of a crawler that collects apps, a repository that stores data, an API gateway for communicating with the client, and a Dapp for communicating with the system nodes on the on-chain. The repository is composed of apps' internal files DB, results DB from the participants in the network, report DB from users, and private data DB that participants need to sell them. The on-chain is built by consortium blockchain. Only participants authorized through the Certificated Authority can participate in the network, and they use the certificates to encrypt data and authenticate themselves. Participants can communicate with its own node by using dapp. The node has a decentralized ledger that records all transactions in the blockchain network. The ledger has the hash of history process that query transaction history, result list, and app address list, etc, and the hash for trading data.

2.2 AI Malware Detection Knowledge-Sharing

The process of detecting Android malware on the proposed platform is carried out in the order of feature extraction and data set generation, detection program AI algorithm learning, input test-set, ranking and setting weights, and repeats learning. Each detection program continuously repeats learning new data to improve the detection accuracy of AI models.

The process of learning data is the same as those of AI methods that previously are studied. After all participants submit each result and the f1-score for test data set, the system ranks the participants based on score. And then, the weight is calculated by weight function that takes the ranking of all participants as input. The result of the application is extracted from the participants' f1-score and ranking weight, and then whether it is a malicious app or not is determined. This labeling for apps help participants process learning data set

The following equations are used to determine whether the application is malicious.

$$m_i = \sum_{j}^{n} w_j \cdot d_{ij} \qquad (1)$$

$$\begin{cases} if\ m_i \le \alpha, & m_i = 0 \\ if\ m_i > \alpha, & m_i = 1 \end{cases}$$

$$\{w\} = W\_RANK(fs_1, fs_2, \dots fs_n) \qquad (2)$$

$$m_i, d_{ij} \subseteq \{0,1\} \qquad (3)$$

Here, $m_i \subseteq \{0,1\}$ is the final label value of the ith application on the test set. $W\_RANK(\,)$ is the weight function that takes the ranking of the participants as input after the ranking is calculated by all participant's f1-score, $d_{ij} \subseteq \{0,1\}$ is the result of the ith application by jth participant. α is the standard score for judging benign and malware of the ith application

3. Experiment

For the first training data, 1000 malware and 1000 benign apps were used, and 50 tests were randomly selected. After detect the test data set, data from previous tests were additionally learned when performing the next test. A total of 5 nodes participate.

The prediction model used was SVM, and it was assumed that all participants used the same SVM model, but each used a different feature model.

Fig. 2 and Table.1 show the comparison of the proposed method and each participant's existing method f1-score. Compared to the data of each participant, it can be seen that the f1-score for detecting malicious apps in the system is a maximum 16% higher.

Fig. 1 shows the result of comparison with the existing method without using the submitted feedback system and the proposed method.

Since the amount of data that was tested and then again was used for the learning step on the next test is smaller than the size of the training data, the initial results are similar to those of the existing method and the proposed method. **Fig. 1** node 3 has the same result until 2450 data. As the test data amount was increased, it was shown that the performance improved in **Fig. 1**

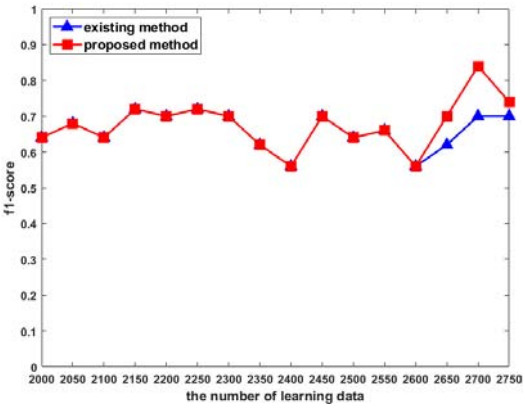


Fig. 1. Comparison with the existing method and the proposed method on node

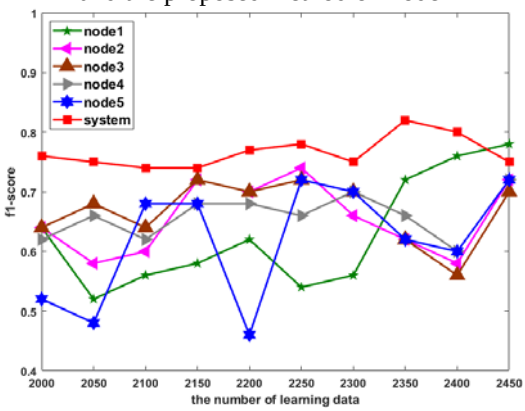


Fig. 2. Comparison with the existing method and the proposed method on all nodes

Table. 1. Node and System f1-score according to the number of learning data

The number of Learning Data	Node & System f1-score					
	1	2	3	4	5	system
2000	0.64	0.64	0.64	0.62	0.52	0.76
2050	0.52	0.58	0.68	0.66	0.48	0.75
2100	0.56	0.60	0.64	0.62	0.68	0.74
2150	0.58	0.72	0.72	0.68	0.68	0.74
2200	0.62	0.70	0.70	0.68	0.46	0.77
2250	0.54	0.74	0.72	0.66	0.72	0.78
2300	0.56	0.66	0.70	0.70	0.70	0.75
2350	0.72	0.62	0.62	0.66	0.62	0.82
2400	0.76	0.58	0.56	0.6	0.6	0.8
2450	0.78	0.72	0.70	0.72	0.72	0.75

4. Conclusions

In this paper, we present a knowledge sharing platform that enables participants to make better knowledge through knowledge sharing. It has been shown to lead to better system outcomes by sharing knowledge created through AI. Not only that, even if knowledge sharing is carried out, participants' valuable personal data must be guaranteed by using the proposed blockchain-based private data market platform. It was shown that the detection performance of malicious apps improved in the proposed platform. However, it was difficult to see that the detection performance of an individual was steadily improved, and it seems to be related to selected the wrong features. Since this paper uses one prediction model, in the future it is necessary to implement the platform through various prediction models.

References

[1] S. Nakamoto. "Bitcoin: A Peer-to-Peer Electronic Cash System" *Online. Available: https://bitcoin.org/bitcoin.pdf*, 2008.

[2] G. Wood, “Ethereum: A secure decentralised generalised transaction ledger,” *Ethereum Project Yellow Paper*, 2014.

[3] S. Rouhani and R. Deters, “Security, performance, and applications of smart contracts: A systematic survey,” *IEEE Access*, vol. 7, pp. 50759–50779, 2019.

[4] X. Lin, J. Li, J. Wu, H. Liang, and W. Yang, “Making knowledge tradable in edge-AI enabled IoT: A consortium blockchain-based efficient and incentive approach,” *IIEEE Transactions on Industrial Informatics*, vol. 15, no. 12, pp. 6367–6378, 2019.

[5] K Liu, S Xu, G Xu, M Zhang, D Sun, H Liu "A Review of Android Malware Detection Approaches Based on Machine Learning," *IEEE Access*, vol. 8, pp. 124579-124607, 2020.

[6] R. Vinayakumar, K. P. Soman, , and P. Poornachandran, “Detecting malicious domain names using deep learning approaches at scale,” *Journal of Intelligent & Fuzzy Systems*, vol. 34, no. 3, pp. 1355–1367, 2018.

SALT Traffic Demand Generation and Calibration for Urban Road Network

Hyewon Song and Moonyoung Chung
Smart Data Research Section
Electronics and Telecommunications Research Institute
Daejeon, South Korea
[e-mail: {hwonsong, mchung}@etri.re.kr]
**Corresponding author: Hyewon Song*

Abstract

The traffic simulation is one of solutions to analyze and forecast the traffic state for road network. In order to simulate the real traffic situation well, the similarity of traffic demands with real traffic demands is important. In this paper, we present the traffic demand generation process for SALT and its calibration process to fit the estimated demand to real demand using diverse traffic data in 4 administrative districts in Gangnam area, Seoul. Finally, compared with measured real traffic volume, we show the estimated demand volume is within about 12% errors coverage.

***Keywords*:** Urban Traffic Simulation, Traffic Demand Generation, Traffic Demand Calibration

1. Introduction

The urban traffic analysis is an essential issue in government strategies, and the traffic simulation technology[1] is one of diverse researches to address the urban traffic state for a large-scale urban area. The SALT[2,3] is a meso-scale traffic simulator for urban road network, and it simulates the large-scale urban area, for example, several administrative districts or one city, with reasonable tradeoff between computational performance and accuracy, compared with a micro-scale simulator. As a key issue for more accurate simulation results, it is important to generate the traffic demand, which is closer to real traffic pattern using limited real traffic data. In this paper, we propose the traffic demand generation process using trajectory data and calibration process using measured volume data such as VDS(Vehicle Detector System) volume data. Also, we apply to 4 administrative districts (including Seocho-gu, Gangnam-gu, Songpa-gu and Gangdong-gu) in Seoul, and then, we show the similarity between estimated results and real traffic data is within reasonable error rates.[4]

2. Weighted Selection based Traffic Demand Generation

In order to reflect real traffic properties to the proposed traffic demand generation (TDG) process (as shown in **Fig. 1**), we propose the TDG weight value matrix derived by statistical properties of measured real traffic data, such as trajectory or VDS data, based on Origin-Destination(OD) pair and time series. Namely, it means the timely volume based pattern of real traffic per OD pair. And then, according to the TDG weight value matrix with given time series and OD pair conditions, we select appropriate trajectories as traffic demands from real trajectory pool. In the process, we use the Flyod-Warsahall Algorithm, which is a kind

This work was supported by Institute of Information & Communication Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2020-0- 00073, Development of Cloud-Edge based City-Traffic Brain Technology).

of the shortest path algorithm, to complementary allocate the demand because of rack of real traffic data. Finally, using SALT, we calibrate the estimated demands repeatedly.

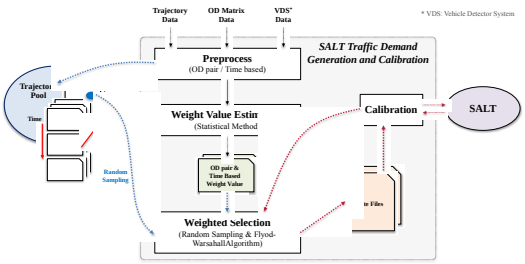


Fig. 1. Conceptual Process Flow of SALT TDG

3. Evaluation and Result

In order to evaluate the similarity of estimated demands with real demands, we use measured volume data for 4 weeks at totally 16 spots as shown in Table 8. Per day of week, we compare the flow volume every 1 hour between estimated data and measured data at the same spot, and we finally show the estimated demand volume is under 12% MAPE(Mean Absolute Percentage Error) averagely.

4. Conclusions

In this paper, we present to generate and calibrate the traffic demands as input data of SALT for 4 administrative districts in Gangnam areas, Seoul, and the estimated demand is under reasonable error rate compared with actual measured volume.

References

[1] S. Pell, et al. “Trends in Real-time Traffic Simulation,” *Transportation Research Procedia*, Vol 25, pp. 1477-1484, 2017.

[2] S. Kim, et al. “SALT-Viz: Real-Time Visualization for Large-Scale Traffic Simulation”, *in Proc. of IEEE Transportation Electrification Conference and Expo,Asia-Pacific (ITEC Asia-Pacific)*, 2019.

[3] H. Song, et al. "Statistical traffic generation methods for urban traffic simulation," in *Proc. of International Conference on Advanced Communication Technology (ICACT)*, 2018.

[4] P. Sekuła, et al. “Estimating historical hourly traffic volumes via machine learning and vehicle probe data: A Maryland case study”, *Transportation Research Part C*, Vol. 97, pp 147-158, 2018.

Table 1. Comparison of Estimation Accuracy at measurement spots

Measurement Spot			*Measured Volume (per 1 day)*	*Estimated Volume (per 1 day)*	*MAPE (%)*
Gangdong	Cheondong E.	S->N	18,336	15,856	11.3
		N->S	30,157	30,335	4.4
	Seongsim H.	E->W	36,232	35,201	6.9
		W->E	33,417	35,201	5.5
Gangnam	Seolleong S.	E->W	42,850	43,902	4.9
		W->E	60,587	62,873	5.4
	Maebong T.	S->N	48,162	49,673	8.6
		N->S	38,706	39,859	6.7
Songpa	Mongchon S.	E->W	20,721	20,050	5.6
		W->E	25,052	24,754	3.5
	Seockchon L.	S->N	32,951	31,591	4.7
		N->S	38,752	34,010	11.2
Seocho	Seocho S.	S->N	75,844	74,380	2.7
		N->S	81,200	81,941	2.1
	Express Bus T.	E->W	67,293	63,485	5.1
		W->E	70,439	68,820	3.1

Adjustment of color coordinates to improve JPEG compression performance

Shahid Abbas, Oh-Jin Kwon and Seungcheol Choi
Department of Electronics Engineering, Sejong University
Seoul, South Korea
[e-mail: shahid@sju.ac.kr, choisc@sju.ac.kr, ojkwon@sejong.ac.kr]
*Corresponding author: Oh-Jin Kwon

Abstract

JPEG is the most well-known standard for image compression and has been widely used for storing and transmitting images. JPEG has been providing quality images with reasonable compression. However, improvement is still appreciated to give better quality with less or the same bitrate. This paper proposes a new approach of color coordinate tuning (CCT) for JPEG image coding. The color coordinate rotates with a best-fit angle to maximize the variance of the Y component. The results show the effectiveness of the proposed method which outperforms the JPEG image coding standard.

***Keywords*:** JPEG, Color Coordinate Tuning, YCbCr color coordinate, Image Coding, Matrix Rotation

1. Introduction

JPEG [1] has been the most widely used image coding standard of lossy compression for storing and transmitting photographic images. It is used for a variety of image applications requiring a compression standard, such as desktop publishing, graphic arts, color facsimile, photo videotex, newspaper wirephoto transmission, and medical imaging owing to its simple structure.

JPEG is a general purpose international standard to meet the needs of almost all continuous-tone still-image applications [2]. Although a considerable number of images distributed over the internet are compressed with JPEG [3], it still has a limitation that large volumes of data such as ITP [4], internet facsimile transmission [5],and DICOM [6] images are transmitted at a relatively lower bandwidth. Image compression presents an interesting yet challenging dilemma for the market, i.e., a tradeoff between the size of the image and image quality. When the compression ratio is increased, the quality of the image reduces, and vice versa. The challenge here, however, is to achieve the best quality with minimum size. JPEG compression ratio is typically in the range of about 10:1 to 20:1 [7].

The primary purpose of the JPEG image coding standard is to store the necessary information and to remove the redundant data from the image while compressing. This standard achieves much of its compression by exploiting the limitation of the human eye. Because the human visual system is less sensitive to the higher frequency, and by suppressing these higher frequency components, more compression is achieved.

Many studies have focused on addressing high compression by using an appropriate color

(a) (b)

Fig. 1. General Framework of JPEG image coding standard: (a) Encoder and (b) Decoder

transformation, which is discussed in Section 2 briefly, such as from RGB to YCbCr. RGB color space is common and used in most computer displays along with television, cell phones, etc. which is non-linear with visual perception and device dependent. Every component of an image in RGB color space is highly correlated, therefore, they are redundant in encoding point of view. On the other hand, In YCbCr color space, Luminance or Y components are more prominent, while Chroma components have much lower entropy hence, they may be encoded more efficiently.

A general framework of JPEG is used to compress either full-color or grayscale images. In the case of color images, RGB is transformed into YCbCr color space to perform different processing on the luminance and chrominance components of the image. After color conversion, initially, DCT is applied on 8x8 blocks of the image. Afterward, quantization is performed on the DCT coefficients. A zigzag scan is carried out for each 8x8 block to exploit high-frequency components and to retain DC and low-frequency components. Then comes the lossless step where Run Length and Huffman coding are used to encode the zigzag-scanned coefficients. The image can be reconstructed through a decompression process, a reversed order of all encoding processes.The overall architecture of the coding system of JPEG is described in **Fig. 1**.

A tradeoff exists between the compression ratio and the quality of the image. The quality of the image reduces as the compression ratio is increased, and vice versa. The challenge here, however, is to achieve maximum or same compression with the best quality image.

This paper proposes a method of color coordinate tuning for the YCbCr color space. The method analyzes and finds the best-fit angle then rotates all the color components to maximize the Y component variance.

The rest of this paper is arranged as follows: Section 2 presents a brief overview about JPEG and related work. Section 3 explains our proposed CCT Algorithm. Section 4 describes the experimental setups and results. Finally, Section 5 concludes the paper and describes the future direction.

2. Related Work

Color images are commonly represented as a combination of different components, for example, red, green, and blue in the RGB color space. When the components are highly correlated, the same information is replicated during the encoding of each of the components. Hence it is suggested to convert to the image in a color space that is as decorrelated as possible. Therefore, most of the images are converted into another color space from RGB, such that YCbCr. For example, in **Fig. 2**, all the three components of the image are more prominent regarding details of the objects in the image (edges, etc.). Compared with **Fig. 3**, however, the Y or luminance component has most of the details about the objects in the image, while the Cb and Cr components have lower entropy or variation. Therefore, the YCbCr color space can be encoded much more efficiently by compression algorithms.

Most work has been done to improve color transform for image compression. Malvar and

(a) (b) (c) (d)

Fig. 2. RGB components : (a) Original Image (b) R-Comp (c) G-Comp and (d) B-Comp

Y-Comp Cb-Comp Cr-Comp

Fig. 3. YCbCr components

Sullivan proposed a new color model [8], called YCoCg-R, as defined below:

$$\begin{bmatrix} Y \\ Co \\ Cg \end{bmatrix} = \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ -1 & 0 & -1 \\ -1/2 & 1 & -1/2 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Hyun Mun Kim *et al.* [9] have proposed a new color model, called YSbSr, which gives high decorrelation gain with small rounding and conversion error as follows:

$$\begin{bmatrix} Y \\ Sb \\ Sr \end{bmatrix} = \begin{bmatrix} 0.6460 & 0.6880 & 0.6660 \\ -1.0 & 0.2120 & 0.7880 \\ -0.3220 & 1.0 & -0.6780 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Roman Starosolski [10] proposed a simple color space transformation for lossless image compression and compared them with existing transformations.

Another color model, called CMY(K), for printing and hard copy output by Adrian Ford [11] is more dependent on the device with a type of inks and paper used, as well as a type of printing device.

Besides, Janak Porwal [12] introduced a new approach to transform a 3-dimensional RGB color space to another cGST 4-dimensional color space, and an additional dimension was introduced for the most real-life images.

However, the above-discussed methods are all about different transforms from RGB color space, while the proposed method adjusts the same YCbCr color space coordinates by rotating.

3. Proposed Method

In JPEG Standard, the first and essential step is to convert the color space of the image from RGB to YCbCr [13]. In YCbCr color space, the Y component shows the brightness, and Cb, Cr component represents the chrominance of the image. This conversion allows gains in terms of the compression by using the characteristics of the human visual system, which is more sensitive to brightness information than color information. [14].

After the conversion, the proposed CCT algorithm is applied on the image by rotating each color coordinate of YCbCr to maximize the variance of Y component.

3.1 Rotation Matrix

Let Y, Cb, and Cr component be in the x , y, and z axes, respectively, in the 3D coordinate [15], and let Y*,Cb*,Cr* are the new color coordinates obtained after rotation with an angle of φ. A 3D rotation matrix rotated around the z-axis, as shown in **Fig. 4**, is obtained as follows:

$$\begin{bmatrix} Y^* \\ Cb^* \\ Cr^* \end{bmatrix} = \begin{bmatrix} cos\varphi & sin\varphi & 0 \\ -sin\varphi & cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} \quad (1)$$

Similarly, A 3D rotation matrix obtained from rotation around x-axis with an angle of rotation θ is

$$\begin{bmatrix} Y^* \\ Cb^* \\ Cr^* \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\theta & sin\theta \\ 0 & -sin\theta & cos\theta \end{bmatrix} \begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} \quad (2)$$

and rotation around y-axis having an angle of rotation $\emptyset$:

$$\begin{bmatrix} Y^* \\ Cb^* \\ Cr^* \end{bmatrix} = \begin{bmatrix} cos\emptyset & 0 & -sin\emptyset \\ 0 & 1 & 0 \\ sin\emptyset & 0 & cos\emptyset \end{bmatrix} \begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} \quad (3)$$

3.2 CCT Algorithm

The proposed CCT consists of the following three steps:

1. Finding variance values of each color coordinate

2. Finding the best-fit rotation angle to maximize the variance for the step 1).
3. Rotate coordinate

In the first step, the mean value for each color coordinate is calculated and subtract that average value from each sample of corresponding channel, making the median to be the center. After that, calculate the variance through the equation given as follows:

$$Y_{var} = \sum_{i=0}^{N} \frac{(Y_i - Y_{mean})^2}{N} \qquad (4)$$

Where Y_{var}, Y_i, and N are the variance, pixel value, and number of pixels respectively.

To maximize the variance in step 2), taking the partial derivative of $Y *= Ycos\varphi + Cbsin\varphi$ from 1). The best-fit angle of rotation is calculated as follows:.

$$\sum(Y^*)^2 = \sum(Ycos\varphi + Cbsin\varphi)^2$$

$$2(Ycos\varphi + Cbsin\varphi)(-Ysin\varphi + Cbcos\varphi) = 0$$

$$tan2\varphi = \frac{2YCb}{Y^2 - Cb^2}$$

$$\varphi = \frac{1}{2} atan \frac{2YCB}{Y^2 - Cb^2} \qquad (5)$$

in Step 3), rotate the color coordinates with the calculated best-fit angle in equation 5).

Similarly, Repeat the above mentioned steps to calculate the best-fit angle for other color channels.

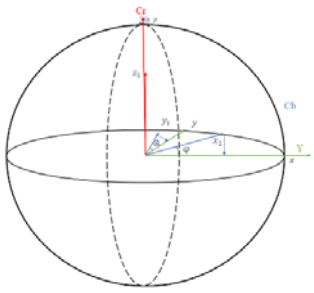


Fig. 4. Rotation about $z - axis$

4. Experiment

To show the effectiveness of the proposed method, we compared the performance with libjpeg-turbo [16], which is a reference software of the JPEG image coding standard using sample images.

4.1 Experimental Setup

After color space transformation from RGB to YCbCr, we applied the proposed CCT on the image to rotate all the color coordinates with the best-fit calculated angle to maximize the variance. **Fig. 5.** shows the sample images [17]. All sample images are compressed with different quality factors, ranging from 50 to 90.

Fig. 5. Sample Images

4.2 Experimental Results

The proposed method was developed with the goal to obtain better quality of images with the same or less bitrate. **Fig. 6.** shows that the proposed method outperformed compared to libjpeg-turbo.

The quality of an image is measured in terms of PSNR against bit per pixel (BPP). Matlab application functions were used to calculate the PSNR of respective image and BPP were calculated through the formula given below:

$$\boldsymbol{BPP} = \frac{Total\ bits}{W * H}$$

Where *Total bits* is the size of image (bytes)*8 and *W* and *H* is the width and height of the image.

5. Conclusion & Future Work

This paper investigates the behavior of color space which shows significant influences with

quality having the same or less bitrate. The proposed CCT method shows that all the color coordinates with a best-fit angle rotated to maximize the variance of Y component provides an improvement of the PSNR performance.

In future we will use this on specific blocks adaptively. The specific block is to be found by using artificial intelligence techniques.

Acknowledgements

This work was supported by the Institue for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No.2017-0-01667, Development of the next generation 360-degree image/video format).

References

[1] Pennebaker, William B and Mitchell, Joan L, " JPEG: Still image data compression standard", *Springer Science \& Business Media*, 1992.

[2] G. Wallace, "The JPEG still picture compression standard", *Communications of the ACM*, vol. 34, no. 4, pp. 30-44, 1991.

[3] "JPEG - JPEG Systems", Jpeg.org, 2020. [Online]. Available: https://jpeg.org/jpegsystems/privacy_security.html.

[4] S. Raman, H. Balakrishnan and M. Srinivasan, "ITP: an image transport protocol for the Internet", IEEE/ACM Transactions on Networking, vol. 10, no. 3, pp. 297-307, 2002.

[5] "Internet Fax Transmission", Support.ricoh.com, 2020. [Online]. Available: http://support.ricoh.com/bb_v1oi/pub_e/oi_view/0001036/0001036377/view/fax/unv/0036.htm.

[6] B. Ramakrishnan and N. Sriraam, "Internet transmission of DICOM images with effective low bandwidth utilization", Digital Signal Processing, vol. 16, no. 6, pp. 825-831, 2006.

[7] Singh, Sukhpal, "Reduction of Blocking Artifacts In JPEG Compressed Image", *arXiv preprint arXiv:1210.1192, 2012.*

[8] Malvar, Henrique and Sullivan, Gary, "YCoCg-R: A color space with RGB reversibility and low dynamic range", *ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q*, vol. 6, 2003.

[9] Kim, Hyun Mun, Woo-Shik Kim, and Dae-Sung Cho. "A new color transform for RGB coding." *In 2004 International Conference on Image Processing*, 200, vol. 1, pp. 107-110, 2004.

[10] R. Starosolski, "New simple and efficient color space transformations for lossless image compression", *Journal of Visual Communication and Image Representation*, vol. 25, no. 5, pp. 1056-1063, 2014.

[11] Ford, Adrian and Roberts, Alan, "Colour space conversions", *Westminster University, London*, vol. 1998, pp. 1-31, 1998.

[12] Porwal, Janak, "A 3D→ 4D color space transform for efficient lossless image compression", *2013 Visual Communications and Image Processing (VCIP)*, pp.1-6, 2013.

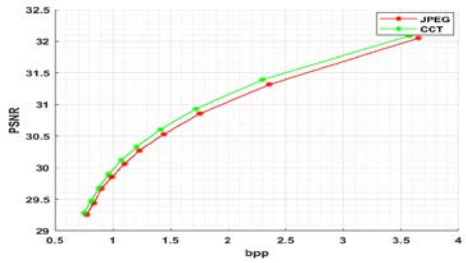
[13] "YCbCr", En.wikipedia.org, 2020. [Online]. Available: https://en.wikipedia.org/wiki/YCbCr.

[14] 2020. [Online]. Available: https://www.sciencedirect.com/topics/engineering/humanvisual-system.

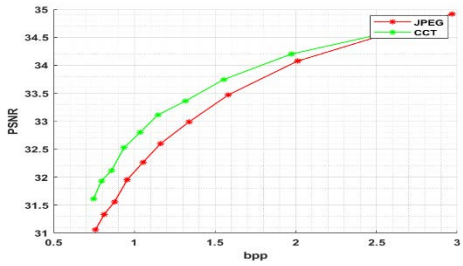
[15] Dunn, Fletcher, Parberry and Ian, "3D math primer for graphics and game development", *Jones \& Bartlett Publishers*, 2010.

[16] "libjpeg-turbo | Main / libjpeg-turbo", Libjpeg-turbo.org, 2020. [Online]. Available: https://libjpeg-turbo.org/.

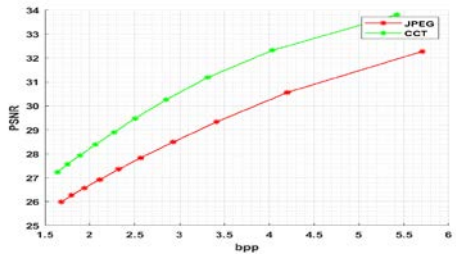
[17] "IMAGES PPM", Igm.univ-mlv.fr, 2020. [Online]. Available: http://igm.univ-mlv.fr/~incerti/IMAGES/PPM.htm.



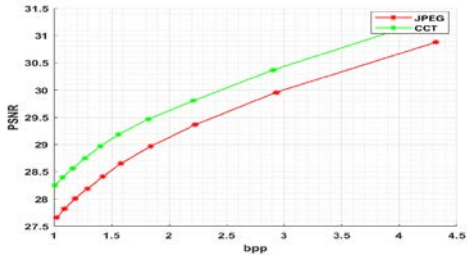
Chili



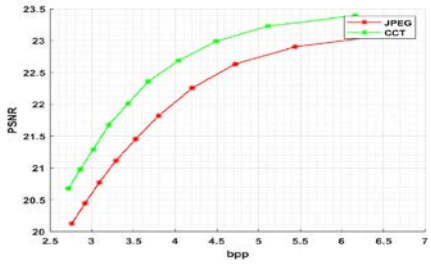
Aeroplan



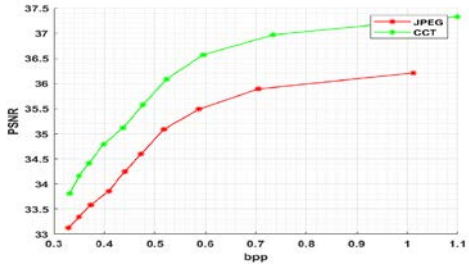
Map



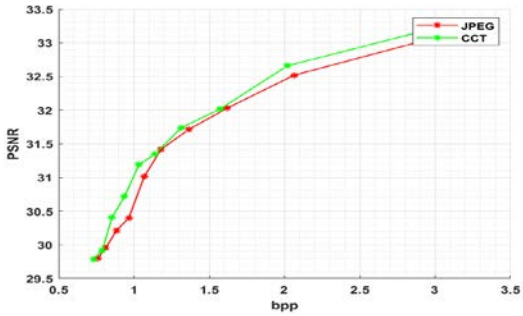
Sailboat



Cartoon



Colors



House

Fig. 6. Results

KSII The 12th International Conference on Internet (ICONI) 2020.

Experience Augmentation for Fast Deep Reinforcement Learning

Mingyu Pi and Moonyoung Chung
Smart Data Research Section
Electronics and Telecommunications Research Institute
Daejeon, South Korea
[e-mail: {mgpi, mchung}@etri.re.kr]
*Corresponding author: Moonyoung Chung

Abstract

The advantage of reinforcement learning is that it learns by itself without human intervention and background knowledge. The RL agent selects the optimal action based on rewards from their own experience and exploration, but because of this characteristic, takes a long time to train. While humans can learn similar experiences through one experience, machines can only learn about the experiences they have acted on. In this paper, we present a method to fast reinforcement learning using experience augmentation. We apply our method to the Cartpole environment and find that training time significantly reduces in both the DDQN(Double DQN) and 3DQN(Dueling Double DQN) algorithms.

***Keywords*:** Reinforcement Learning, Experience Augmentation, Deep Q-Networks

1. Introduction

Reinforcement learning is one area of machine learning in which agents solve problems by self-learning without human intervention or background knowledge. With this advantage, attempts to apply reinforcement learning to solve problems in various domains are increasing. However, it takes a long time to learn because it learns based on the agent's actions and rewards without background knowledge. This is because humans can learn similar things together through one experience, while machines learn only one experience. To solve this problem, we propose a method called experience augmentation(EA). The data augmentation[1] is already widely used in the image classification problem. This method augments the training dataset by rotating, flipping, or cropping an image to increase performance. In this paper, we conduct experiments to augment the agent's experience based on this idea and find that the training time reduces in the Cartpole environment [2].

2. Experiment Augmentation

We propose a method for reinforcement learning agents to learn similar experiences simultaneously through one experience, like to humans. Each time the agent acts, it adds a similar experience to the replay memory along with the experience. Flipping, rotation, adding noise, etc. can be a way to create this similar experience. However, it may be difficult to apply depending on the environment. For example, in the Cartpole environment, if the pole goes below the horizon, the game ends, so there is only a value for the environment above the horizon. So it is difficult to use rotation to augment the experience. Therefore, in the next section, we use flipping for the experiment.

This work was supported by Institute of Information & Communication Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2020-0- 00073, Development of Cloud-Edge based City-Traffic Brain Technology).

Table 1. Compare the number of episodes performed to the end of the training

Measures	*Method*	*Mean*	*SD*	*t-statistics*	*p-value*
The number of episodes performed until the end of the training	DDQN	155.0	31.3	6.2003	3.296e-05***
	DDQN + EA	86.9	15.0		
	3DQN	172.0	39.8	5.4827	3.523e-05***
	3DQN + EA	**80.5**	34.7		
The number of episodes that last 500 steps for the first time	DDQN	120.6	24.4	6.0451	1.808e-05***
	DDQN + EA	64.4	16.4		
	3DQN	121.9	25.6	6.5116	4.966e-06***
	3DQN + EA	**54.2**	20.7		

***p<0.001, N=10

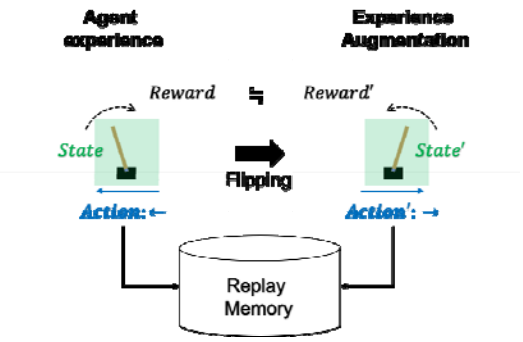


Fig. 1. Example of experience augmentation in the Cartpole environment

3. Experiment

In this section, we conduct experiments using DDQN and 3DQN algorithms. We show quantitative evaluation and t-test results by comparing to the traditional reinforcement learning method and show that reduces training time.

3.1 Experiment setting

We use Double DQN(DDQN) [3] algorithm and Dueling Double DQN(3DQN) algorithm (Dueling DQN [4] + Double DQN) for the experiment in the Cartpole environment. In one episode, the game ends when the pole of Cartpole goes down the horizon or if the pole does not go down for 500 steps. Training ends when the agent lasts 500 steps for 5 consecutive episodes. Since DQN uses a neural net, there may be differences in training time and results depending on the initial weight value, so it performs 10 times for each algorithm.

3.2 Experiment result

We use two measures to compare the difference in training time between applying and not using our proposed method. The first is the number of episodes performed until the end of the training, and the second is the number of episodes that last 500 steps for the first time. As shown in Table 1, our method has fewer episodes performed until the end of the training, and we verify whether the results are significant by performing a t-test.

4. Conclusion

In this paper, we propose a method to fast reinforcement learning using experience augmentation. Experience augmentation methods include flipping, rotation, cropping, and adding noise. We apply one of the EA method using DDQN and 3DQN algorithm in the Cartpole environment and find that the training time significantly reduces.

References

[1] Perez, Luis, and Jason Wang. "The effectiveness of data augmentation in image classification using deep learning." *arXiv preprint arXiv*:1712.04621, 2017.

[2] https://gym.openai.com/envs/CartPole-v1/

[3] Van Hasselt, Hado, Arthur Guez, and David Silver. "Deep reinforcement learning with double q-learning." *arXiv preprint arXiv*:1509.06461, 2015.

[4] Wang, Ziyu, et al. "Dueling network architectures for deep reinforcement learning." in *Proc. of International conference on machine learning (ICML)*, 2016.

KSII The 12th International Conference on Internet (ICONI) 2020.

An Efficient Method for Ground-truth Generation of Image Forgery Datasets

Nam Pham[1], Sihyun Park[2], Chunsu Park[1, 2*]
[1] AI Pro Inc., Seoul, South Korea
[2] Department of Computer Education, Sungkyunkwan University, Seoul, South Korea
[e-mail: cspk@skku.edu]
*Corresponding author: Chunsu Park

Abstract

CASIA v1 and v2 are well-known datasets and have been widely used in image forgery detection and localization research. Nevertheless, due to the lack of ground-truth images, previous studies in the field have difficulty in benchmarking their performances when using these datasets. In this paper, we present an efficient method to generate the ground-truth images for CASIA datasets with pixel level precision.

Keywords: ***dataset; forgery; generation; ground-truth***

1. Introduction

Image forgery detection and localization (IFDL) has been actively studied in recent years. CASIA Image Tampering Detection Evaluation (ITDE) [1] are phenomenally popular datasets, which have been widely utilized to evaluate the researches in IFDL field. Nevertheless, the authors of CASIA ITDE datasets do not publish the ground-truth images which encouters difficulty that there is no standard criteria for researchers to benchmark their localization performance.

In this paper, we propose an efficient method to generate ground-truth images for CASIA datasets and make it publicly available online at https://github.com/namtpham/casia2ground-truth

2. CASIA Forgery Datasets

Columbia datasets [2] were known as the first datasets in IFDL studies with gray (v1) and color images (v2). The main disadvantages of Columbia Datasets are their simplicity. As a result, CASIA have gained widespread popularity in the field thanks to their enormous variety and complicated manipulation using Photoshop. Therefore, CASIA tampered images are realistic and challenging to forgery detection methods. CASIA v1 and v2 have 1.725 and 12.323 images, respectively, where the latter is an extesion of the former. Both datasets include 3 types of images: authentic images, spliced images, and copy-move images. The authentic images in these datasets are classified into different categories, such as animal, architecture, character, nature, scene, etc. From the authentic images taken by camera, image regions are copied from one image and pasted into another image or itself to create spliced or copy-move images, respectively. The copied regions may be scaled and/or rotated in the target images to make them deceptively look realistic.

3. Proposed Method

3.1 Ground-truth Image Generation

The tampered images in CASIA datasets are named based on the authentic images which were

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2019R1F1A1055593).

used to create them. For each tampered image T, we can easily extract its corresponding authentic image I. We then apply image differencing between couples of images using multiple values of threshold φ to generate the binary difference images. Due to the compression or post processing effects of the tampered images, the generated images may contain salt and pepper noises. Median filters with different kernel sizes k are applied to remove these noises. Finally, we obtain a set of output generated ground-truth images for each input tampered image. The generated ground-truth image G can be formulated as follows:

$$G = M_k * \left(I - T - \Phi\right),$$

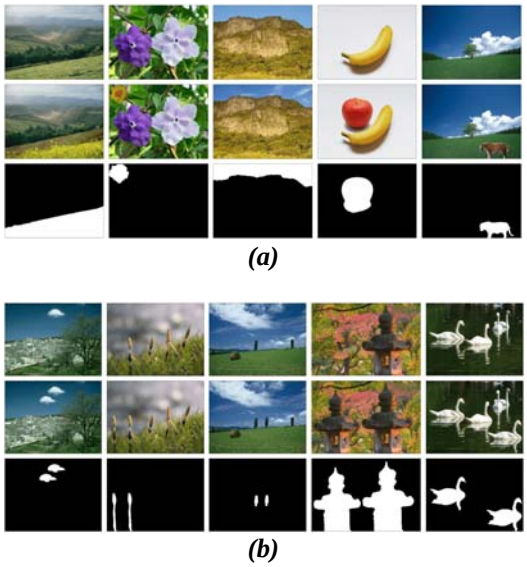
where Φ is the matrix with the same size as I with all elements φ, M_k is median kernel size k, and $*$ denotes the convolution operator. From the generated images, we can manually choose the most appropriate ground-truth image for each tampered image with considerable ease.

Nonetheless, the published datasets have some problems in naming the files, which result in the incorrect generation of multiple ground-truth images. We completely address these problems by finding the matching images using SIFT features.

3.2 Validation of the Generated Images

To validate the accuracy of generated ground-truth images, we adopt the forgery classification and localization method [3-4] where SIFT features and image retrieval are used for two objectives: *(i)* Find relevant images of tampered images; *(ii)* Search for the copy/source regions of copy-move images (the generation process in subsection 3.1 can only localize paste/target regions).

We repeat the process in subsection 3.1 to create the ground-truth images to verify whether the output images in 2 steps are similar. We can also use SIFT features to search for the matching regions of generated masks in the copy-move images. **Fig. 1** ***(a)*** and ***(b)*** show examples of ground-truth images for spliced and copy-move images, respectively. The ratio of number of tampered images after the performance of this process increases from 95.90% to 99.58%.



(a)

(b)

Fig. 1. Examples of ground-truth for ***(a)*** spliced images and ***(b)*** copy-move images generated by the proposed method. In each column, the authentic, tampered, and ground-truth images are in top, middle, and bottom rows, respectively.

4. Conclusions

In this paper, we present an efficient method to generate ground-truth images for popular forgery datasets CASIA. The proposed method can automatically generate ground-truth for 99.58% number of images for both datasets.

References

[1] J. Dong, W. Tang, T. Tan, "CASIA Image Tampering Detection Evaluation Database," in *Proc. of IEEE China Summit and International Conference on Signal and Information Processing*, 2013.

[2] TT. Ng, SF. Chang, "A Dataset of Authentic and Spliced Images Blocks," in *ADVENT Technical Report 203-2004-3 Columbia University*, June 2004.

[3] NT. Pham, JW. Lee, CS. Park, "Structural Correlation Based Method for Image Forgery Classification and Localization," *Applied Sciences*, Vol 10. No. 13, pp. 4458, 2020.

[4] NT. Pham, JW. Lee, GR. Kwon, CS. Park, "Hybrid Image-Retrieval Method for Image-Splicing Validation," *Symmetry, Vol 11. No. 1, pp. 83, 2019.*

A Study on Control Platform Based on Arduino Uno

Dongeon Yoon and Hyosang Lee and Amsuk Oh
Dept of Digital Media Engineering, Tongmyong University
Busan, Korea
[e-mail: ehddjs97@naver.com, gytkd9575@naver.com, asoh@tu.ac.kr]
*Corresponding author: Amsuk Oh

Abstract

This paper presents a platform that can efficiently control elevator door waiting time by using Arduino UNO. The proposed platform provides a way to maximize the convenience of elevator users by controlling elevator doors with short intervals between open and closed. For this operation, an elevator model is produce using Arduino board and elevator layer sensing sensor, elevator door control sensor, step motor and Bluetooth module (HC-06). After that, the company will develop Arduino firmware that will receive and process control command from Android platforms by utilizing the Arduino Integrated Development Environment (IDE). Then implement an Android platform that carries out the transmission of the elevator door control command to perform a Bluetooth connection with the Arduino firmware, then test the elevator model. Since the platform is wirelessly operated in a mobile environment, it can not only provide convenience to users but also ensure safety against diseases in the COVID-19 era.

***Keywords*:** Arduino Uno, Bluetooth, Control platform, Elevator door

1. Introduction

Arduino as a hardware maker tool for efficient elevator door control, and Android[1] as a software maker tool is useful in terms of compatibility. Code designed in the Arduino Integrated Development Environment is designed to work if the values entered on the Android platform meet the conditions, so the elevator doors can be controlled by connecting only Bluetooth. It also operates an ultrasonic sensor via serial communication to maintain a safe distance between the passenger and elevator doors. This prevents passengers from having to wait for the elevator door to open or press the existing elevator button directly with their hands[2].

2. Arduino firmware design

Arduino is a generic term for all development tools and environments related to microcontroller boards that are completed based on open source. Generally, Arduino performs serial communication as shown in Figure 1. Although serial communication is slower than parallel communication, it has the advantage of being able to communicate between devices at low power and low cost. The RX and TX lines send and receive data from each other, and the GND adjusts the amount of power received consistently.

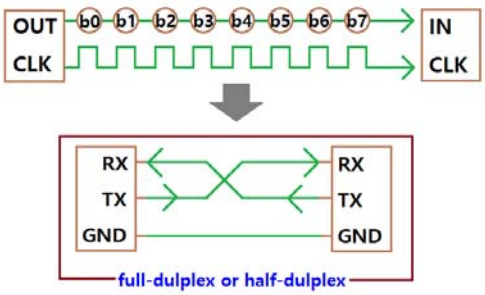


Fig. 1. Arduino serial communication process

In this paper, Bluetooth communication is carried out using HC-06 module as shown in Figure 2, with Arduino hardware playing Slave, and Android platform, a type of mobile application, playing the role of Master.

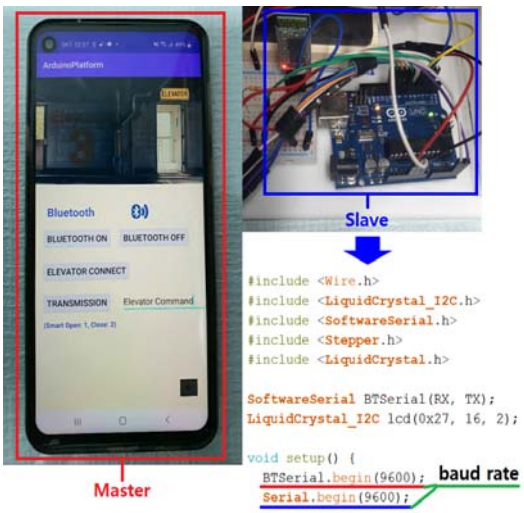


Fig. 2. Relationship Between Arduino and Android Platform

3. Conclusions

In this paper, a method of producing elevator models using Arduino Uno and implementing Android platform to control the waiting time for elevator doors on Bluetooth was presented. Through the design of Arduino firmware, I understood the principles of Arduino serial communication and Bluetooth communication. By implementing the Android platform, we looked at the operating functions of each method. The principle of the step motor operating the elevator door was identified and the actual implementation result was shown. This device prevents the elevator door from closing in front of your eyes and moving to another location. And it is relatively safe because it is not a direct way of pressing the elevator button.

Acknowledgements

This work was supported by the BB21+ Project in 2020.

References

[1] H. Xu, and G. J. Zhao, "Wireless Monitoring System for Elevator on Android Platform," *In Mechatronics and Automatic Control Systems*, pp. 373-381, 2014.

[2] Y. Zhang, X. Sun, X. Zhao, and W. Su, "Elevator ride comfort monitoring and evaluation using smartphones," *Mechanical Systems and Signal Processing*, Vol 105, pp. 377-390, 2018.

[3] Y. W. Lin, and C. Y. Lin, "An interactive real-time locating system based on bluetooth low-energy beacon network," *Sensors*, Vol 18, No. 5, pp. 16-37, 2018.

[4] I. Pérez, J. Á. H. Cuevas, and J. T. E. Galindo, "Design and construction of a desktop AC susceptometer using an Arduino and a Bluetooth for serial interface," *European Journal of Physics*, Vol 39, No. 3, pp. 035203, 2018.

[5] A. S. Lai, and S. Y. Leung, "Mobile bluetooth-based game development using arduino on android platform," *In Applied Mechanics and Materials*, Vol. 427, pp. 2192-2196, 2013.

[6] E. Rice, "Increasing step motor torque without increasing frame size," *Machine design*, Vol 92, No. 4, pp. 22-23, 2020.

[7] M. Yu, B. Xu, and X. Ouyang, "Optical Switch Based on Stepper Motor," *CHINESE JOURNAL OF LASERS*, Vol 46, No. 8, pp.0804003, 2019.

A Study on Implementation of Computer Programming Language by Flowgorithms

Kalaphath Kounlaxay[1], Shinjin Kang[2], Yoonsik Shim[3], and Soo Kyun Kim[4*]
[1]Department of Computer Engineering, Souphanouvong University, Louangprabang, Laos
[2]School of Games, Hongik University, Sejong, Korea
[3]Department of Game Engineering, Paichai University, Daejeon, South Korea
[4]Department of Computer Engineering, Jeju National University, South Korea
[e-mail: kimsk@jejunu.ac.kr]
*Corresponding author: Soo Kyun Kim

Abstract

Computer Programming requires the design of algorithms, which are then written into a computer programming language for processing. Even the small program is syntactic and semantically complex that far exceeds the level of the beginner students. Therefore, this research will introduce how Flowgorithm to convert flowcharts or algorithm to be converted to several real-world programming languages. Proposed flowcharts can be executed directly by a built-in interpreter, that allows students to learn programming concepts before confronting the specific syntax and semantic of the language.

***Keywords*:** Algorithm, Flowgorithm, Programing Language, Source Code Viewer, Console

1. Introduction

Flowgorithm is a program that allows you to create a flowchart very easily. Usually, program is written using a text editor depending on the programming language. How to teach algorithms and programming as part of computational thinking is still an open question [1]. By using a flowgorithm, it uses shapes that represent different actions, to understand algorithms, rather than conventional programming languages that need to understand the structure of each language. There are many flowchart-based programming environments for improving the comprehension of the problem-solving skills of novice programmers that can be found in [2]. We can run programs directly in a flow algorithm, and can also convert them to commands in almost 20 languages such as C++, C#, Pascal, Java, JavaScript, Perl, Python, MathLab, PHP, Visual Basic .NET etc. [3]. Moreover, there is an e-book created by Roberto Atzori with more than 250 flowcharts. Flowgorithm was created by Devin Cook (the last version 2.18.3 is from November 2018).

2. Flowgorithm Features

The classic method of interaction between a computer and a user is by using the "console" when using this method, the user will enter data by using the keyboard, and the program will display output on a text-only screen. Sometimes the console permits text to alter color, but most of them are white text on a black background. Rather than the use of the text-only screen, the algorithm tries to make it seem like a regular instantaneous messenger interface. From a programmer's point of view, it looks like they are sending messages with computers, and the "Chat bubbles" are color-coded to match the input and output shapes used in flowcharts. In the screenshot, the program's output bubble is in green while the user's input bubble is in blue [4]. Flowgorithm has the following features:

- Easily understand output
- Graphics variable monitoring window
- Interactively generate real code
- Safe recursion

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No.NRF-2019R1A2C1002525)

- Loops, arrays, and flexible expressions
- Support multiple languages

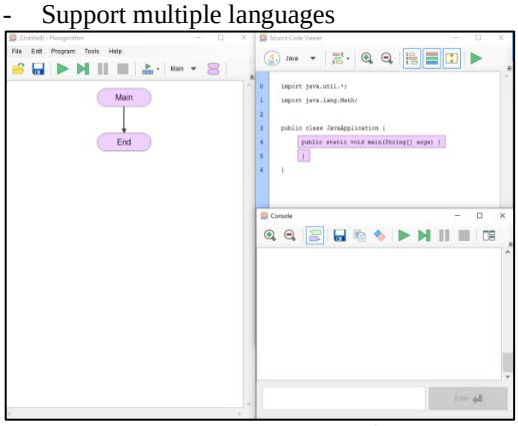


Fig. 1. Flowgorithm Interface

Fig. 1. shows the features of the Flowgorithm interface which consists of three main parts such as Flow-Chart or Algorithm text-based, Source Code Viewer and Console.

3. Implementations

In this paper, we have proposed to use Flowgorithm or flow-chart as algorithm to solve the problems which we can input as text-based programming language then We also can convert them to commands in Source Code Viewer that enables flowcharts to be converted into real-world programming languages in which the generated source code was created by using the program templates. Shown as **Fig. 2.**

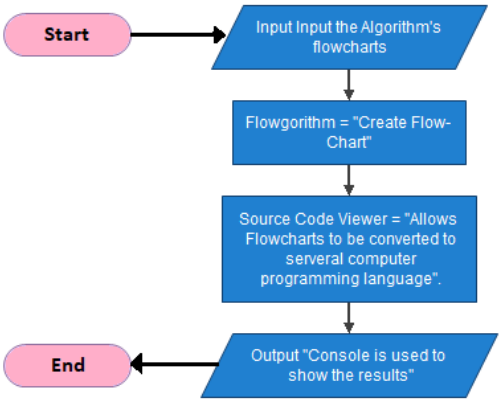


Fig. 2. Flowgorithm Implementation Phases

According to **Fig. 2.** then we have implemented by designed the operation of the program for ATM transaction in case of withdrawing money from an ATM by drawing a flowchart as a sequence diagram in which Flowgorithm's source code viewer that used is in Java language.

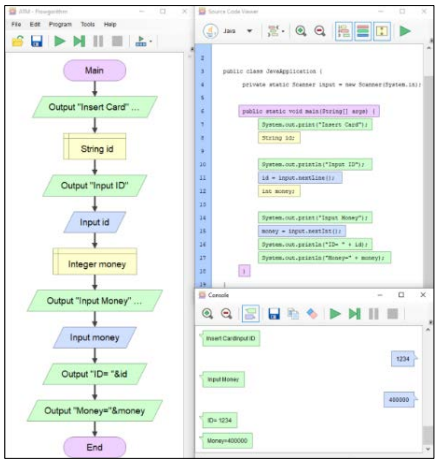


Fig. 3. Results

4. Conclusions

Flowgorithm enabled to help the student show how a shape relates to its generated source code; both shape and source code are color-coded based on their functional category.

Acknowledgement. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No.NRF-2019R1A2C1002525)

References

[1] Wolfram, S. (2016). "How to Teach Computational Thinking". Stephen Wolfram Blog. Retrieved from http://blog.stephenwolfram.com/2016/09/how-to-teachcomputational-thinking.

[2] Hooshyar, D., Ahmad, R. B., Nasir, M. H. N. M., Shamshirband, S., & Horng, S. J. (2015). "Flowchart-based programming environments for improving comprehension and problem-solving skill of novice programmers: a survey". International Journal of Advanced Intelligence Paradigms, 7(1), 24– 56.

[3] Devin D. Cook. (2015). "Flowgorithm: Priciple for Teaching Introductory Programming Using Flowcharts". Proceedings of the 2015 American Society for Engineering Education/Pacific South West Conference.

[4] http://www.flowgorithm.org/documentation/index.htm.

Development of portable plantar pressure mat sensor and monitoring system

Dae Gyeom Kim[1], Young Kim[2], Jun Ma[1], Lyudmila Dmitrievna Khan[1], Min Hong[3] and Se Dong Min[1,4,*]
[1]Department of Software Convergence, Soonchunhyang University
[2]Institute of Wellness Convergence Technology, Soonchunhyang University
[3]Depart of Computer Software Engineering, Soonchunhyang University
[4]Department of Medical IT Engineering, Soonchunhyang University
[e-mail: daegyeom09@sch.ac.kr, ykim02@sch.ac.kr, ringring369@gmail.com, dmitrievna.km@gmail.com, mhong@sch.ac.kr, sedongmin@sch.ac.kr]
*Corresponding author: Se Dong Min

Abstract

Plantar pressure distribution represents a variety of information including the functions of the foot and balancing ability. This study aimed to develop a low-cost, portable plantar pressure mat sensor with a monitoring application for both static and dynamic locomotion measurement. A pilot testing to examine the performance of the developed sensor was performed as a comparison with a commercial plantar pressure measurement system. This study developed a 35 x 50 sensor matrix based on Velostat film for plantar pressure distribution measurements and an application for real-time analysis of lower limb movements. For performance test of the developed sensor, footprint area and the trajectory of center of pressure (COP) were compared with that of MatScan. Pilot data of a male subject (bodyweight: 70 Kg, height: 181 cm, shoe size: 275 mm) were collected for 10 seconds of static standing on both feet and 5 steps of dynamic in-place gait. Comparison results of the developed sensor and MatScan showed an average error of 2.35(±1.86) cm^2 for footprint area, X-axis of COP had an average error of 0.24(±0.34) index, and Y-axis of COP had an average error of 0.92(±1.05) index. The preliminary findings of this development study may be the basis for further development of the sensor. Our future studies will focus on the sensor's feasibility for general use in daily living and expansion to clinical uses by including larger number of subjects in various environment setting.

***Keywords*:** Plantar Pressure, Center of Pressure Trajectory, Footprint Area, Portable, Mat Sensor

1. Introduction

Plantar pressure (PP) is the force that the sole of the foot presses against the ground and its distribution represents a variety of information including the functions of the foot, balancing ability, and risk of falling [1,2]. For easier and more practical measurement of plantar pressure distribution (PPD), many researchers have endeavoured to develop various forms and methods including shoe insoles and exercise mats [3]. One of the popular and research-verified mat-type pressure sensors is force plate, which is known to accurately measure and analyze the key parameters related to human body movements and locomotions, especially PPD and center of pressure (COP) trajectories.

This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Korean government (MSIT) (No. NRF-2019M3E5D1A02069073).
This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2018R1D1A1B07050037).

The demerits of force plate are its high cost and heavy weight, making it difficult to be carried around for outdoor use [4].

Velostat (Adafruit, USA), on the other hand, is a light-weight, flexible material made of polymeric foil (polyolefins) that is electrically conductive. Any pressure applied on Velostat changes resistance as the applying force causes conductive particles to touch, increasing the current through the sensors [1]. Due to its flexibility and low cost, Velostat has been used by many researchers for measurement and analysis of human body posture and foot pressure [1,5,6].

For more various forms of future locomotive researches, this study developed a low-cost portable plantar pressure measurement mat sensor using Velostat film and a real-time monitoring application.

2. Method

2.1 Plantar pressure mat sensor (PPMS)

This study developed a portable plantar pressure mat sensor (PPMS) using Velostat in a 35 x 50 sensor matrix in the size of 0.5 cm^2. Each sensor was located 0.3 cm apart to prevent short circuit. To measure the resistance value of Velostat, copper tape was used for conductor plate. The upper and lower layers of the plate were arranged orthogonally and the point of intersection corresponded to the pressure sensing area. The pressure sensing region was based on the voltage distribution law and the offset resistance was set to 100 Ω. The structure of PPMS is as shown in **Fig. 1** and the total number of pressure detecting area is 1,750.

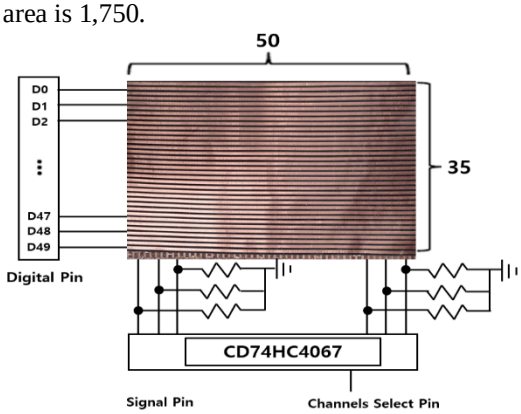


Fig. 1. Structure of PPMS

2.2 Real-time plantar pressure monitoring application

For real-time monitoring and measurement of 1,750 plantar pressure data collected by PPMS, a C# based application was developed. As shown in the upper left corner of **Fig. 2**, the application has a scroll-down window for the users to select the type of ports, connections between PPMS and UART, and data saving functions. Underneath the port configuration window, the grid area in **Fig. 2** presents the plantar pressure data for real-time monitoring in gray scale.

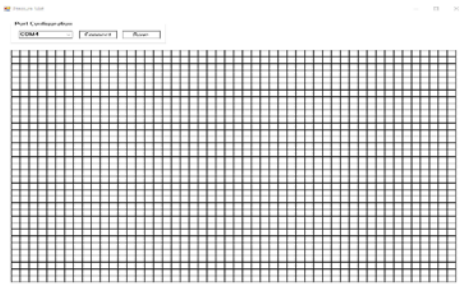


Fig. 2. Real-time monitoring application

2.3 Protocol

For performance comparison analysis of the developed PPMS, a commercial plantar sensor (MatScan) was used. Pressure data from the two sensors were collected simultaneously; PPMS was placed over MatScan. Sampling rates for PPMS and MatScan were both set at 70 Hz. The subject for this study was a 27-year-old male with the bodyweight of 70 Kg, height of 181 cm, and shoe size of 275 mm. For the performance testing of PPMS, the subject was to stand still for 10 seconds and then to take 5 consecutive steps, respectively for static and dynamic movement analysis based on plantar pressure distribution. This experiment was performed to detect and calculate Center of Pressure (COP) Trajectory and Footprint Area, where the bodyweight was acting as pressure against the sensor. The subject was instructed to step on PPMS and MatScan on his third step among the 5 steps for the stability of the data [7].

2.4 Data processing and analysis

Data collected in the static standing position were used to calculate the footprint area to which plantar pressure (bodyweight) was applied. For accurate data analysis, the first and last 3 seconds of the collected data were deleted and only the

mid 4 seconds of data were used. The data collected from consecutive 5 steps were used for COP calculation by using Equation 1 [8].

$$CoP_x = \frac{\sum_{i=1}^{n} X_i \times P_i}{\sum_{i=1}^{n} P_i} \quad CoP_y = \frac{\sum_{i=1}^{n} Y_i \times P_i}{\sum_{i=1}^{n} P_i} \quad (1)$$

3. Results

Footprint area comparison between PPMS and MatScan showed an average error of 2.35(±1.86) cm2. **Fig. 3** shows the PPD comparison result of PPMS and MatScan.

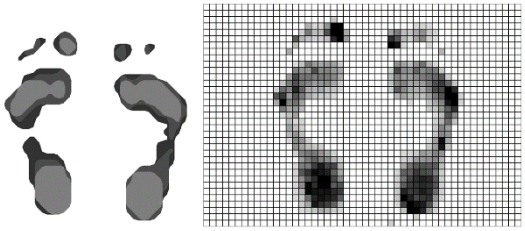


Fig. 3. PPD comparison result of PPMS and MatScan

COP comparison results of PPMS and MatScan showed an average error of 0.24(±0.34) index on the X-axis of COP and an average error of 0.92(±1.05) index on the Y-axis.

For COP trajectory comparison, the number of channels was used as the index. The X-axis had 1 to 50 indices, and the Y-axis had 1 to 35 indices. **Fig. 4** shows the COP trajectory comparison result of the PPMS and MatScan.

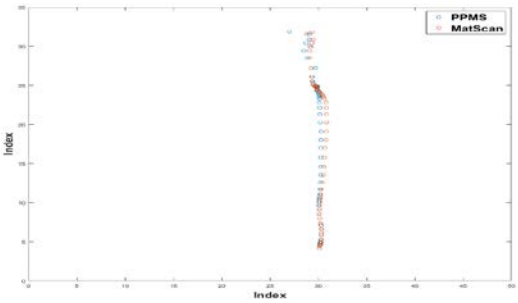


Fig. 4. COP Trajectory comparison result of PPMS and MatScan

4. Conclusions

This study developed a low-cost, portable plantar pressure mat sensor based on Velostat together with a real-time monitoring application for both static and dynamic human locomotion measurement. The preliminary findings of this study may be the basis for further development and use of the sensor.

References

[1] SUPRAPTO, S. S., et al. Low-cost pressure sensor matrix using velostat. In: 2017 5th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME). IEEE, 2017. p. 137-140.

[2] CHESNIN, Kenneth J.; SELBY-SILVERSTEIN, Lisa; BESSER, Marcus P. Comparison of an in-shoe pressure measurement device to a force plate: concurrent validity of center of pressure measurements. Gait & posture, 2000, 12.2: 128-133.

[3] TAN, Adin Ming, et al. Design of low cost smart insole for real time measurement of plantar pressure. Procedia Technol, 2015, 20.6: 117-122.

[4] LEE, Cha Hyuk, et al. Development of a Foot Plantar Pressure Sensing System and Characteristic Analysis of the Sensing Signal along Gate Status. In: Institute of Control, Robotics and Systems Domestic Conference Paper. 2011. p. 654-659.

[5] JEONG, Eunseok; LEE, Jaehong; KIM, DaeEun. Finger-gesture recognition glove using velostat (ICCAS 2011). In: 2011 11th International Conference on Control, Automation and Systems. IEEE, 2011. p. 206-210.

[6] VEHEC, Igor; LIVOVSKY, Lubomir. Flexible Resistive Sensor Based on Velostat. In: 2020 43rd International Spring Seminar on Electronics Technology (ISSE). IEEE, 2020. p. 1-6.

[7] ZAMMIT, Gerard V.; MENZ, Hylton B.; MUNTEANU, Shannon E. Reliability of the TekScan MatScan® system for the measurement of plantar forces and pressures during barefoot level walking in healthy adults. Journal of foot and ankle research, 2010, 3.1: 11.

[8] AQUEVEQUE, Pablo, et al. Gait Segmentation Method Using a Plantar Pressure Measurement System with Custom-Made Capacitive Sensors. Sensors, 2020, 20.3: 656.

Design of Untact Distance Learning Platform based on OTT devices

Yejin Jang[1], Hajeong Yoo[2], Yumin Jo[3], Jongho Paik[1] and Siyoung Lee[4]
Department of {Software Convergence[1], Digital Media Design and Applications[2], Mathematics[3]},
Seoul Women's University and Aloys Inc[4]
Seoul and Seongnam, South Korea
[e-mail: {yuvm02, hjyoo0813, rosemin97, paikjh}@swu.ac.kr, andy@aloys.co.kr]
*Corresponding author: Jongho Paik

Abstract

As people are isolated due to the Covid-19, self-isolation and social distance has become an big issue. Due to the reasons for refraining from going outside, the new term 'untact' has advent. Untact means contactless, not meeting people in person. In the field of education, particularly, real-time online education is provided along with existing online education. In this paper, we propose the design of an untact online education service platform for distance learning based on OTT devices. Also, we propose a MVC pattern to develop proposed service platform development.

***Keywords*:** Untact, contactless, Distance learning, Education contents, OTT devices

1. Introduction

Nowadays, due to the Covid-19 pandemic, many students are affected by school closures. Globally, more than 1.3 billion students are unable to attend school[1]. In South Korea, while the school closure, the Department of Education facilitate school to educate students at home via digital devices such as computer or TV. As a result, education has changed dramatically with the rise of online education learning, using digital platforms.

Currently, most forms of online courses are taking on smart devices or computer. However, there are several problems relate to inconvenience. On top of that, students suffer from accessing the network through smart devices. Furthermore, students struggle to listen classes for a long periods of time in small screen.

Moreover, user interface and serveral steps to access online platform is too complex. Therefore, in this study, we propose the design of untact distance learning service platform based on OTT devices in MVC pattern.

2. Advanced Research

2.1 OTT(Over-the-Top) device

OTT stands for "over-the-top," the term used for the delivery of film and TV content via the internet, without requiring users to subscribe to a traditional cable or satellite pay-TV service like a Comcast or Time Warner Cable[2].

As Korean content-specialized video sites such as Viki.com and Drama Fever have grown in the U.S. and global markets, they can be easily consumed through new media platforms such as OTT. In particular, in a media environment where online mobile content consumption is rapidly increasing, the new media platform is expected to become an even more important role[3].

2.2 Real-Time Online Education Platform

Minerva School in the U.S. offers classes through the real-time video-based 'Active

Learning Forum'[4]. This platform has several functions such as group discussion, 1:1 feedback and student management. In addition, Korean cyber universities such as Seoul Cyber University and Hanyang Cyber University have communication between professors and students through video lectures and chatting easily.

3. Design of Untact Distance Learning Platform

As shown as **Fig. 1**, the overall structure of this service platform is presented. In this paper, we design a platform that can support online education in OTT devices. To build Online Education Platform peer-to-peer, we need to merge OTT broadcast service and server based on MVC pattern.

After installing the OTT devices, the platform is designed with personalized UI/UX for students to use it comfortably at home. The platform's server is designed as Node.js in that Education should be conducted in real-time. Node.js is an asynchronous I/O-based architecture that is suitable as a real-time server capable of parallelizing processes[5].Therefore, the platform's UI/UX uses the template EJS and PUG for Node.js. Also, this service platform is required a relationship between tables, MySQL is suitable as it is flexible and highly scalable. In addition, it is needed to store photos and videos, so we prefer to use Firebase, which is also available for Google Analytics as system administrator's dashboard. It is shown in (**Fig. 2**).

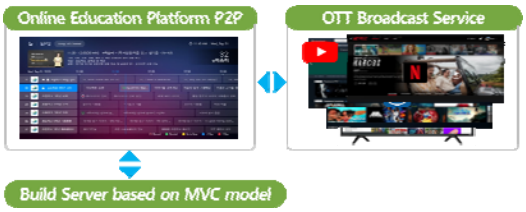


Fig. 1. Structure of Untact Distance Learning Platform based on OTT devices

3. Conclusions

In this paper, we proposed the design of untact distance learning platform based on OTT devices. To make this platform in peer-to-peer, we suggested merging OTT broadcast service and

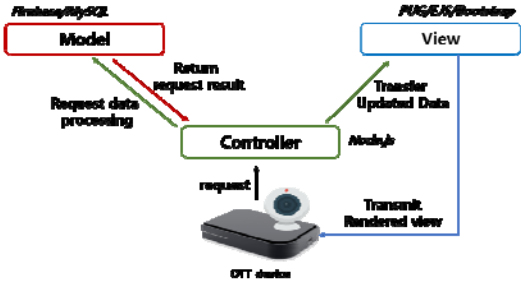


Fig. 2. MVC Pattern of Platform

server which is based on MVC pattern. We proposed this platform to be designed in MVC pattern. In Model, we preferred to use MySQL and Firebase. For MySQL, it is possible to make a relationship between tables and it also functions high flexibility and scalability. For Firebase, it can not only save photos and videos, but it is suitable to use Google Analytics as a dashboard for system administrator. In View, we proposed to use EJS and PUG template adequate for using Node.js. Lastly, in Controller, we suggested to use Node.js as this platform operates in real-time. Through this paper, we expect the enhancement and improvement of untact distance learning platform.

Acknowledgment

This work was supported by the Technology development Program(S2967732) funded by the Ministry of SMEs and Startups(MSS, Korea).

References

[1] https://en.unesco.org/covid19/educationresponse/solutions

[2] https://digiday.com/media/what-is-over-the-top-ott/

[3] YoungHwan Kim, HoeKyung Jung, "Business Model of New Media Platform in K-Content Use," Journal of Digital Convergence, Vol 14, No.10, pp. 431-438, 2016.

[4] https://www.kird.re.kr/repository/uploadfiles/board/1581550608097.pdf

[5] Hajeong Yoo, Yumin Jo, Yejin Jang, Jongho Paik, "A Patrol Route Guidance Recommendation Service Platform based on User' Desire Information," in *KSII the 15th Asia Pacific International Conference on Information Science and Technology (APIC-IST)*, 2020.

The Effect of Quality of Education System and Teachers' Characteristics on Educational Performance: Focusing on the Moderating Role of Online Educational Fitness

Xiu Jin, Chunfan Jin, and Sangwoo Hahm
Department of Business Administration, Honam University
Gwangju, South Korea
Normal College, Yanbian University
Yanji, China
Department of Business Administration, Semyung University
Chungbuk, South Korea
[e-mail: soohua1120@gmail.com, yitswjcb@163.com, thehahm@semyung.ac.kr]
Corresponding author: Sangwoo Hahm

Abstract

Both organizations and businesses are being negatively affected by the coronavirus. In particular, most schools that can be regarded as one organization are implementing online education. In general, students took offline classes, but the coronavirus caused a change in taking online classes. As a result, we believe that students' educational performance can be affected. With regard to online classes, students can use mobile devices to take classes at home and elsewhere, not at school. This study aims to measure the students' educational performance due to online classes. As the characteristics of online education, it focused on the system quality of school online education and teachers' characteristics. Therefore, measuring educational performance according to the level of all variables is the most importanct purpose. In addition, it focused on the suitability of online education for students as a way to enhance the impact of the system quality of education and teahcers' characteristics on the educational performance. Therefore, it is intended to verify what kind of changes can be brought about the educational performance through the moderating effect of the online education fit by classifying the groups with high and low levels of online education fit. Through empirical analysis, the quality of the online education system will improve students' educational performance. And the teachers' characteristics will also improve students' educational performance. Lastly, the higher the student's fit for online education, the greater the quality of the online education system and the teachers' characteristics will have a greater impact on educational performance. This study would like to present the measures necessary to increase the outcomes of students' online education. In addition, it would like to discuss various variables that can play a key role in improving performance. Finally, it would like to suggest directions for future research in relation to online education.

Keywords*:* Mobile device, Online system, Quality of education system, Performance

1. Introduction

Today, the novel Coronavirus is currently affecting every aspect of our daily lives around the world [1]. Due to the coronavirus, most institutions are offering online classes [2]. In this regard, the quality of online education may impact on students' educational performance. Each school has its own online education system. So it can be seen that these quality of systems are different. This difference can be seen as a key variable that determines the level of students' educational performance. Furthermore, it can be seen that the teachers' characteristics also have an effect on educational outcomes. In terms of characteristics, it focuses on the teacher's online teaching ability and the online lectures efficacy. Thus, when teachers have strong online lecture skills, the quality of knowledge delivered to students will increase, and furthermore, there is a high possibility of improving academic performance. In addition, the efficacy of online lectures can be seen as the belief that teachers can successfully conduct online lectures. So, the higher the efficacy of online lectures, the greater the likelihood of showing their ability. This can play a role in enhancing student educational performance. Furthermore, this study focuses on the moderating effect of online lecture fitness in this relationship. Therefore, this study aims to verify this with a focus on deriving a plan to improve students' educational performance.

2. Literature Review

formal education system means the skills and knowledge provided within premises of institutions such as colleges and universities [3]. In this study, the teachers' characteristics are classified into the teachers' online lecture ability and online lecture efficacy. Online lecture ability is defined as the ability that teachers can effectively deliver instructional positions to students through online lectures. The teacher's online lecture efficacy means as the belief that teachers can lecture well through a mobile device. Academic performance is defined as the percentage of aggregate marks that students attained university examination [4]. According to this concepts, educational outcomes are defined as outcomes related to students' overall academic performance. Online education fitness is defined as the overall degree of suitability related to efficiently educating students using mobile devices.

Hypothesis 1: The quality of the online education system will have a positive influence on educational performance.

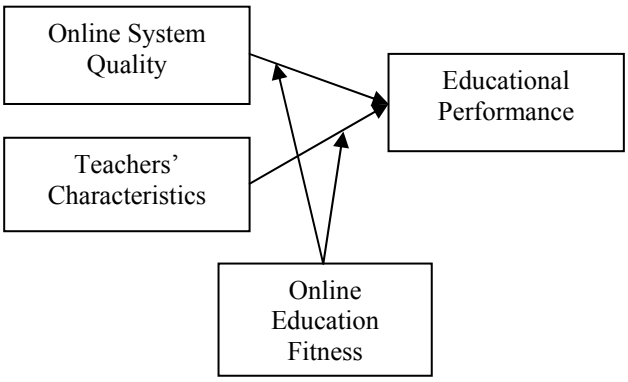
Hypothesis 2: Teachers' online lecture will have a positive influence on educational performance.

Hypothesis 3: Teachers' online lecture efficacy will have a positive influence on educational performance.

Hypothesis 4: Online education fitness will moderate the relationship between teachers' online lecture ability and educational performance.

Hypothesis 5: Online education fitness will moderate the relationship between teachers' online lecture efficacy and educational performance.

3. Research Model



4. Results

The quality of the online education system can enhance educational performance. Teachers' online lecture can enhance educational performance. Teachers' online lecture efficacy can enhance educational performance. Online education fitness can moderate the relationship between teachers' online lecture ability and educational performance. Online education fitness can moderate the relationship between

teachers' online lecture efficacy and educational performance.

References

[1] A. A. Fernandez and G. P. Shaw, "Academic Leadership in a Time of Crisis: The Coronavirus and COVID□19," Journal of Leadership Studies, Vol 14, No. 1, pp. 39-45, 2020.

[2] W. Bao, "COVID‐19 and online teaching in higher education: A case study of Peking University," Human Behavior and Emerging Technologies, Vol 2, No. 2, pp. 113-115, 2020.

[3] P. Kalimasi, "Assessment of entrepreneurship education trends in the formal education system in Tanzania," Business Management Review, Vol 21, No. 2, pp. 53-65, 2018.

[4] S. S. Vinakoti, G. R. Damaji, L. Kasala, N. K. Zakkula, and L. Maturi, "Effect of sleep on academic performance among male and female medical students," National Journal of Physiology, Pharmacy and Pharmacology, Vol 9, No. 12, pp. 1180-1184, 2019.

An Energy-efficient Control Approach for Configuration Propagation with Self-triggered Control in Multi-agent Networking Systems

Yamin Han and Heejung Byun
Department of Information & Telecommunication Engineering, Suwon University
Gyeonggi-do, South Korea
[e-mail: heejungbyun@suwon.ac.kr]
*Corresponding author: Heejung Byun

Abstract

Multi-agent networking systems (MNSs) exhibit complex emergent behavior by using local control laws based on spatial information from nearby environment and adjacent agents. The challenge in the MNS design is developing techniques for the operator to interact with the MNS in order to make system behavior adaptive to changes in system configuration and for operator commands without having to handle them individually. Another challenge is saving energy consumption over the agents in the MNS by reducing the number of information exchange between agents when the system configuration is spread out the network from the operator. To address these challenges, this paper presents an energy-efficient control approach for indirect system configuration propagation with self-triggering control. The proposed method controls the MNS operation by indirectly propagating the system configuration within the framework of local rules. Moreover, a self-triggered propagation model is designed according to the convergence rate of configuration propagation in order to save and to balance energy consumption among agents in the MNS. Theoretical analysis and simulation results are performed to demonstrate the superiority of the proposed method.

***Keywords*:** Wireless sensor networks, energy-efficiency, multi-agent systems.

1. Introduction

A multi-agent networking system (MNS) consists of a large number of agents that coordinate autonomously based on underlying control laws [1]-[3]. Given that the consensus behavior generated is collective and scalable, as well as robust to failure of an individual agent, MNSs are useful in complex tasks including environmental exploration, large-scale search and rescue, and protection [4]. In an MNS, an operator directs agents to carry out mission goals or tasks [5]-[12]. The resulting behaviors the MNS generates depend on a set of parameters of agent algorithms or choice of system parameters for their operation. The types of control an operator can exert on the agents are the following: switching between algorithms that implement desired consensus behavior, changing parameters of algorithms, controlling through selected agent members, remote programming, new software downloading, and reprogramming [5]-[6], [21].

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2020R1A2C1004390).

As such, in order to perform supervisory control of consensus behaviors, the central design challenge in MNS is developing techniques for the operator to convey appropriate parameter adjustments or system configuration independent of the number of agents as intended goals change [13]-[15]. Specifically, when the user input is sequentially applied to the MNS, the operator needs to estimate the optimal time to allot for the next input to the system. This is because system performance is affected by the time between control inputs that the operator applies to the system [16]. However, operators have difficulty understanding the evolution of the system. Another design challenge is saving the energy consumption over the agents in MNS. The agent members in the MNS are generally battery-powered, so they can be easily depleted of energy if they remain active while these controls take place in the MNS. Consequently, energy will be imbalanced among agents and MNS will have shorter lifetime. Furthermore, the operator is not aware of these local battery states of agents and do not identify the number of agents the system configuration has been propagated to. Thus, energy is wasted because of the continuous spread of such controls to the system.

Walker et al. [17]-[18] focused on two methods of information propagation (flooding and consensus methods) and compared the ability of operators to manage the MNS to the desired goals. In the flooding method, each agent explicitly matches the value of user command. Meanwhile, in the consensus method, each agent matches the average value of user command of all the neighbors it senses. They also investigated the use of dynamically selected leaders that are directly controlled by the operator to guide the rest of the MNS [18]. Goodrich et al. [19] worked on a leader-based control of systems using tele-operated leaders based on Couzin's control laws. Pendleton et al. [20] similarly implemented a leader-based model using both virtual agents and an operator as leaders in a system. McLurkin et al. [21] proposed a single-hop broadcast algorithm for downloading new software, but agents that are too far from the user will not be reprogrammed. Li et al. [22] proposed an architecture for multi-agent communication networks, in which agents are clustered to one or multiple systems and each system can be monitored by some central servers through a wireless mesh backbone. Chen et al. [23] proposed a generic framework for the multi-agent planning solution, i.e., the determination of the number of agents. Dmarogonas et al. [24] proposed event-driven strategies to reduce the number of the control updates. Each agent computes its next update time and performs a self-triggered setup. The aforementioned existing studies addressed the interaction problems between the operator and the multi-agent system, but limited work has focused on how system configuration should be spread through the MNS via operator-agent interactions while achieving the system's energy efficiency. We previously proposed a simple model of configuration propagation by a human operator [25]. Such model allowed each agent to automatically drive the system parameters to the desired configuration without considering the system's energy efficiency.

This paper extends the previous work and proposes a dual control approach for indirect propagation of system configuration with an energy efficient self-triggering control in MNSs. First, we propose a method that influences MNS operation by indirectly propagating system configuration from the operator within the framework of local rules in the MNS. Second, we design a self- triggering propagation model of agent state, in which each agent autonomously determines when to send its configuration state update to the neighbors depending on the configuration propagation rate. Then, we extend the self-triggering model to an optimal timing control, where the operator computes the optimal time to give sequential control input to the MNS. Finally, based on theoretical analysis, we provide insights into the performance of the proposed method by deriving the convergence to the desired goal and the stability of the proposed system.

2. Proposed Method

2.1 Configuration State Model

We consider an MNS consisting of N agents and an operator. Let $\boldsymbol{N} = \{1,2,...,N\}$ denote the set of agents in the MNS. The set of neighboring agents of agent i $(1 \leq i \leq N)$ is denoted as Ni $(\subseteq \boldsymbol{N})$, and its cardinality is N_i. The N-th agent in N is referred to as a gateway agent. In this section, a system configuration propagation model is proposed. In

the model, the operator controls the MNS by conveying configuration setting to the single gateway agent, then influences indirect propagation to the rest of the MNS based on autonomous control laws each agent uses. No control can be exerted over individual agent members, only over the MNS as a whole.

For the configuration state control, a discrete-time formula is developed in which the control updates take place every control period τ_s. Thus, the time is divided into $[t, t+1)$, $t=0,1,\ldots$, with time duration equal to τ_s. First, $u(t) \in \boldsymbol{R}^n$ is denoted as the desired system configuration issued by the operator at time slot t, such as command inputs, parameter setting, and parameter changes that the operator wants to achieve with the MNS. Let the vector of the configuration states of all agents at time slot t as or $\boldsymbol{x}(t) = [x_1(t)\ x_2(t)\ \ldots\ x_N(t)]^T$, where $x_i(t) = [x_i^1(t)\ x_i^2(t)\ldots\ x_i^n(t)]^T \in \boldsymbol{R}^n$ is the configuration state vector of agent i at time slot t and individual agent has n-dimensional configuration state space.

The operator interacts with the MNS by applying the desired configuration input to the gateway agent, while the other agents control their respective configuration state vectors and propagate them by interacting with each other. A simple configuration model has been introduced in the previous work [25]. Specifically, the previous work used the following local law of agent i ($i \in \boldsymbol{N}\backslash N$) in the MNS:

$$x_i(t+1) = x_i(t) + \alpha \sum_{j \in \mathbf{N}_i \backslash N} (x_j(t) - x_i(t)) + \alpha 1_{(i,N)}(x_N(t) - x_i(t)),$$

where, α is the parameter that determines the convergence speed and $1_{(i,N)}$ is the indication function: the value of $1_{(i,N)}$=1, if node N (the gateway) is a neighbor of node i and $1_{(i,N)}$=0, otherwise. The operator gives commands or informs the gateway agent of the desired configuration sets and parameter changes with $u(t)$, after which the user input is transformed into configuration state vector and propagated to the MNS. To achieve this, the user input is first applied for the single gateway agent by letting $x_N(t)=u(t)$, then the interactions between the remaining agent members are handled autonomously. Specifically, each agent in the MNS interacts with its neighbors with configuration state vector x_i rather than actual values of system configuration u. That is, the user's intended command is conveyed only to the gateway agent and each of the remaining agents performs indirect parameter setting aided by the proposed dynamics instead directly propagating the user command. In this way, the operator takes actions independent of the number of agents and supervises the MNS as a single entity, hence a control complexity of $O(1)$.

2.2 Self-triggered Propagation Model

When the system configuration issued by the operator is propagated throughout the MNS, an important aspect to consider is to save energy consumption by reducing the number of messages to be exchanged while keeping the energy consumption of agent members balanced. In this section, a self-triggered propagation model is designed for each agent based on the configuration states. The self-triggering state of agent i ($\in \boldsymbol{N}$) is denoted as r_i, which is determined by the following rule:

$$r_i(t+1) = r_i(t) - \epsilon_r(r_i(t) - r_{\min}) + \frac{\left\| \sum_{j \in \mathbf{N}_i} (x_j(t) - x_i(t)) \right\|}{\left\| \sum_{j \in \mathbf{N}_i} x_j(t) \right\|},$$

where ϵ_r is the control parameter to be chosen and $r_{\min}$ is the minimum message triggering probability that can be set by the operator. When agent i calculates r_i, it independently generates a random value ω following the uniform distribution within [0, 1]. If the value of r_i is greater than ω, the agent triggers transmission of its configuration state update to the neighbors. The higher the value of r_i, the higher the probability of message transmission. The self-triggered propagation model works in such a way that the minimum message triggering probability is achieved when agent i reaches an equilibrium point, where the state of the agent's configuration converges to the user command u. As the configuration state of agent i is closer to the averaged state among its neighbor agents, the value of r_i goes near to $\underline{r_{min}}$. The configuration state of an agent becomes identical to those of its neighbors, and the value of r_i decreases; hence, the lower the probability of message exchanges and energy savings. On the contrary, as the configuration state of agent i is different from those of surrounding neighbors, the value of r_i increases, resulting in a higher message triggering probability and frequent packet transmission. This accelerates the convergence

speed of the configuration state values to u for all agents in the MNS.

2.3 Optimal Control of Sequent Input

Another important function of the MNSs is to estimate the system state so the operator can change or properly give sequence control input to the MNS. To support this feature, we present an optimal timing control for sequence input. Based on the self-triggering state control, the gateway agent determines the optimal time to give the next input based on the value of r_N as follows:

$$r_N(t+1) = r_i(t) - \epsilon_r(r_N(t) - r_{\min}) + \frac{\left\| \sum_{j \in N_N} (x_j(t) - x_N(t)) \right\|}{\left\| \sum_{j \in N_N} x_j(t) \right\|}.$$

The gateway agent provides a feedback of r_N value to the operator. The foregoing stability analysis indicates that when the states of all agents converge to the desired configuration, the systems is then driven to $r_N = r_{\min}$. Therefore, the operator can understand system states and estimate the timing to give for the next control inputs, considering the feedback from the gateway agent only. Define $r_d = r_{\min} - r_N$. Then, the operator gives a new input to the MNS when $r_d < \varepsilon$, where ε is the tolerance for convergence.

3. Simulation Results

To show the performance of energy efficiency and balance, we set the initial power (E_0) for each differently. Figure 1 (a) shows the energy difference between agents with the proposed scheme and with the consensus scheme. The consensus method has no appropriate mechanisms for operation adaptation to schedule agent's state, so the agents are activated stochastically. Each agent randomly generates a number and compares it with a given threshold. If the probability value is larger than the threshold, the agent becomes activated. In this simulation, the threshold values are set to 0.3 (mid-th), 0.02 (min-th), and 0.7 (max-th), respectively. For the metric of energy difference ratio, the smaller the energy difference is, the more the energy balance is achieved. The proposed method shows that the energy difference ratio decreases over time and goes near zero after 350 min. Therefore, energy

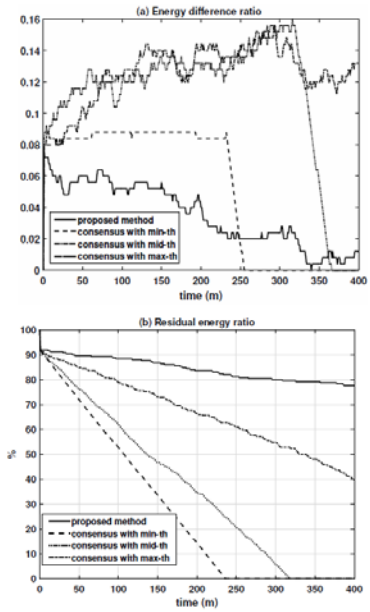


Fig. 1. Energy efficiency: (a) energy difference ratio between the agent with the highest energy consumption and that with the lowest one in the MNS and (b) residual energy ratio in the MNS.

consumption ratio is balanced among the agents in the MNS. However, in the consensus method, the energy difference ratio is maintained or escalates. In the plot of the consensus method with min-th, the energy difference ratio suddenly drops to zero after 250 min, thus the agents' batteries are exhausted. The energy depletion with min-th occurs within a shorter period as compared to other threshold values. This is caused by more frequent activation of agent. **Fig. 1 (b)** shows the residual energy ratio of the proposed and the consensus schemes. Result suggests the proposed method incurs the highest available energy ratio; whereas, in the consensus method, energy depletion occurs relatively early, hence a reduction in network lifetime. The consensus method with min-th also shows a much shorter network lifetime than with other threshold values. In short, the proposed method automatically adjusts the operation state of each agent so energy consumption ratio is balanced among the agents. The proposed method can likewise avoid unnecessary energy waste and increase network lifetime effectively by automatically adjusting the operation state according to the configuration state.

Fig. 2 (a) shows the effect of varying the value of r_{min} on the residual energy ratio. Such an increase of r_{min} reduces the residual energy

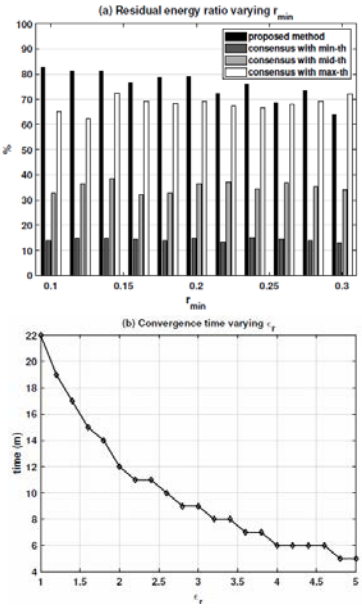


Fig. 2. . Effects on control parameters: (a) residual energy ratio varying r_{min} and (b) convergence time varying ε_r.

ratio, hence a shorter network lifetime. In other words, if the value of r_{min} is increased, although the steady state is entered, the active period is extended, resulting in unnecessary energy waste. Specifically, when the value of r_{min} increases to 0.3, the residual energy ratio of the consensus method becomes somewhat higher than the proposed method. Therefore, the appropriate value of r_{min} needs to be determined and adjusted for future works. **Fig. 2 (b)** shows the convergence time with different values of ε_r. Convergence time is defined as the time when the difference between the value of operation state (r_i) and r_{min} falls below 0.01. The convergence time decreases as the value of ε_r increases. However, for a smaller value of ε_r, the operation state exactly converges to r_{min}; whereas, for higher value of ε_r, the operation value continues to oscillate around r_{min}. Hence, choosing the optimal value of ε_r should also be studied in the future.

4. Conclusions

This paper presents a dual control approach for energy-efficient MNS interaction system. First, for the system configuration control, the proposed scheme indirectly controls the consensus operation of the MNS by propagating the configuration state values to the MNS based on the proposed control laws of each agent. Second, we propose a controller for the agent's operational state scheduling according to the configuration propagation rate. The proposed algorithm forwards the following major contributions. First, each agent automatically drives the system parameters to the desired system configuration even in an erroneous environment. Second, each agent effectively controls its operation mode according to the configuration state, thus balancing energy consumption in the MNS. Finally, insights into the theoretical analysis of the proposed scheme are provided by deriving the system convergence and proving the system stability.

References

[1] M. Brambilla, E. Ferrante, M. Birattari, and M. Dorigo, "Swarm robotics: a review from the swarm engineering perspective," *Swarm Intell.*, vol. 7, no. 1, pp. 1-41, 2013.

[2] Y. Tan and Z. Y. Zheng, "Research advance in swarm robotics," *Defence Technol.*, vol. 9, no. 1, pp. 18-39, 2013.

[3] A. Liekna and J. Grundspenkis, "Towards practical application of swarm robotics: Overview of swarm tasks," *Engineering for Rural Development*, pp. 271-277, 2014.

[4] A. Kolling, P. Walker, N. Chakraborty, K. Sycara, and M. Lewis, "Human interaction with robot swarms: A survey," *IEEE Trans. on Human-Machine Systems*, vol. 46, no. 1, pp. 9-26, 2016.

[5] Z. Kira and M. A. Potter, "Exerting human control over decentralized robot swarms," *Proc. Int. Conf. on Autonomous robots and agents*, pp. 566-571, 2009.

[6] M. Lewis, "Human interaction with multiple remote robots," *Rev. Human Factors Ergonomics*, vol. 9, no. 1, pp. 131-174, 2013.

[7] S. Kerman, D. Brown, and M. Goodrich, "Supporting human interaction with robust robot swarms," *Int. Conf. Resilient Control Systems*, pp. 197-202. 2012.

[8] J.-P. de la Croix and M. Egerstedt, "A control Lyapunov function approach to human-swarm interactions," *Proc. Amer. Control Conf.*, pp. 4368-4373, 2015.

[9] M. A. Goodrich, B. Pendleton, and S. Kerman, "What types of interactions do bio-inspired robot swarms and flocks afford a human?," *Proc. Robotics, Science, and Systems*, pp. 105-112, 2012.

[10] J. Gancet, E. Motard, A. Naghsh, C. Roast, M. M. Arancon, and L. Marques, "User interfaces for human robot interactions with a swarm of robots in support to firefighters," *Int. Conf. Robotics and Automation*, 2010.

[11] S. Bashyal and G. K. Venayagamoorthy, "Human swarm interaction for radiation source se- arch and localization," *Proc. IEEE Swarm Intelligence Symposium*, pp. 1-8, 2008.

[12] B. Tabibian, M. Lewis, C. Lebiere, N. Chakraborty, K. Sycara, S. Bennati, and M. Oishi, "Towards a cognitively-based analytic model of human control of swarms," *Proc. AAAI Spring Symp. Series*, pp. 68-73, 2014.

[13] S. Nagavalli, L. Luo, N. Chakraborty, and K. Sycara, "Neglect benevolence in human control of robotic swarm," *Int. Conf. Robotics and Automation*, 2014.

[14] K. Sycara, C. Lebiere, Y. Pei, D. Morrison, Y. Tang, and M. Lewis, "Abstraction of analyti- cal models from cognitive models of human control of robotic swarms" *Int. Conf. Cognitive Model*, pp. 13-18, 2015.

[15] P. Walker, S. Nunnally, M. Lewis, A. Kolling, N. Chakraborty, and K. Sycara, "Neglect bene- volence in human control of swarms in the presence of latency," *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 3009?3014, 2012.

[16] P. Walker, S. A. Amraii, M. Lewis, N. Chakraborty, and K. Sycara "Human control of leader- based swarms," *Proc. IEEE Int. Conf. Syst. Man, Cybern.*, pp. 2712-2717, 2013.

[17] P. Walker. S. A. Amraii, N. Chakraborty, and K. Sycara, "Human control of robot swarms with dynamic leaders," *Proc. IEEE/RSJ Int. Conf. Intell. Robot*. Syst., pp. 1108-1113, 2014.

[18] M. Goodrich, B. Pendleton, P. Sujit, and J. Pinto, "Toward human interaction with bio-inspired robot teams," *Proc. IEEE Int. Conf. Syst. Man, Cybern.*, pp. 2859-2864, 2011.

[19] B. Pendleton and M. Goodrich, "Scalable human interaction with robotic swarms," *American Institute of Aeronautics and Astronautics*, 2015 (http://dx.doi.org/10.2514/6.2013-473).

[20] J. McLurkin, J. Smith, J. Frankel, D. Sotkowitz, D. Blau, and B. Schmidt, "Speaking Swarmish: Human-Robot Interface Design for Large Swarms of Autonomous Mobile Robots", *AAAI Spring Symposium: To Boldly Go Where No Human-Robot Team Has Gone Before*, pp. 72-75. 2006.

[21] M. Li, K, Lu, H. Zhu, M. Chen, S. Mao, and B. Prabhakaran "Robot swarm communication networks: Architectures, protocols, and applications," *Int. Conf. Communications and Networking in China*, 2008.

[22] M. Chen, L. T. Yang, T. Kwon, L. Zhou, and M. Jo, "Itinerary Planning for Energy-Efficient Agent Communications in Wireless Sensor Networks," *IEEE Trans. on Vehicular Technology*, vol. 60, no. 7, pp. 3290-3299, 2011.

[23] D. V. Dmarogonas, E. Frazzoli, and K. H. Johansson, "Distributed Event-Triggered Control for Multi-Agent Systems," *IEEE Trans. on Automatic Control*, vol. 57, no. 5, pp. 1291-1297, 2012.

[24] H. Byun, "Configuration Dissemination with State Feedback for Wireless Sensor Networking Systems," *Sensors and Materials*, vol. 31, no.59, pp. 1545-1555, 2019.

[25] M. Mesbahi and M. Egerstedt, "Graph theoretic methods in multi-agent networks," *Princeton University Press*, 2010.

[26] C. C. Moallemi and B. V. Roy, "Consensus propagation," *IEEE Trans. on Information Theory*, vol. 52, no. 11, pp. 4753-4766, 2006.

KSII The 12th International Conference on Internet (ICONI) 2020.

Performance Evaluation of Container-Based Clouds with Kubernetes

Younghwan Shin, Byounghoon Son, Jeongyoon Shin, Sangdo Kim, and Jong-Moon Chung
School of Electrical and Electronic Engineering, Yonsei University
Seoul, South Korea
[e-mail: shinyh1115@yonsei.ac.kr, sonhun99@yonsei.ac.kr, sjy970528@yonsei.ac.kr, sangdokim96@yonsei.ac.kr, jmc@yonsei.ac.kr]
*Corresponding author: Jong-Moon Chung

Abstract

As computationally intensive applications such as blockchain became pervasive, most cloud computing platforms are struggling with load balancing problem. Kubernetes and Docker Swarm are no exception. Kubernetes and Docker Swarm are a container-based cloud platform that manages operation of container deployment and scaling in a clustered environment. In this paper, the load balancing performance of Kubernetes and Docker Swarm is provided in terms of throughput and response time.

***Keywords*:** Blockchain, Kubernetes, Docker Swarm, Cloud Computing, Container.

1. Introduction

As complex and heavy applications such as blockchain and machine learning are flourishing, container-based cloud platforms including Kubernetes have widely been adopted. Container-based cloud platforms are lighter, smaller, and faster compared to existing virtual machine platforms [1]. The main reason is that deploying a full operating system instance is utterly unnecessary in container-based cloud platform. There are many container-based cloud platforms. Examples include Kubernetes and Docker Swarm.

However, not much has been written in comparing container-based cloud platforms, which makes it difficult to decide which platform to deploy. Some argumnents including [2] have been made that Kubernetes is the most efficient platform among many container ochestrators but experimental results are still insufficient.

The aim of this paper is to provide comparison of load balancing performance between Kubernetes and Docker swarm. The two platforms adopted in this paper are arguably well known.

2. Simulation Settings

To benchmark the load balancing performance, a word counting web program was built, and deployed on a local computer, Minikube-based local Kubernetes sever, Google Cloud-based remote Kubernetes server, and Docker Swarm-based local server. A local computer without any virtualization was chosen as the baseline since it would ultimately have the best performance.

Each server was consistently given a 16.7kB text file by users. This workload arrival follows Poisson process with 10 ms of mean. The word counting web program could offer up to 30 threads to users. Thus, maximum 30 users can simultaneously use the program. The transaction throughput and response time were measured. The transaction throughput means the number of text file processed per 1 second.

3. Simulation Results

In **Fig. 1**, the higher transaction throughput indicates higher performance. The transaction throughput increases when the number of users is

small. However, after the number of users exceeds the specific point, the transaction throughput decreases and is maintained at lower values. This phenomenon occurs generally among all the cases. This is because all of the platforms suffered the load balancing problem.

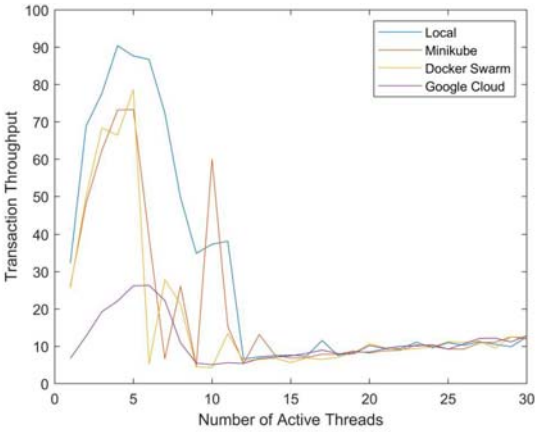


Fig. 1. Transaction throughput based on the number of active threads

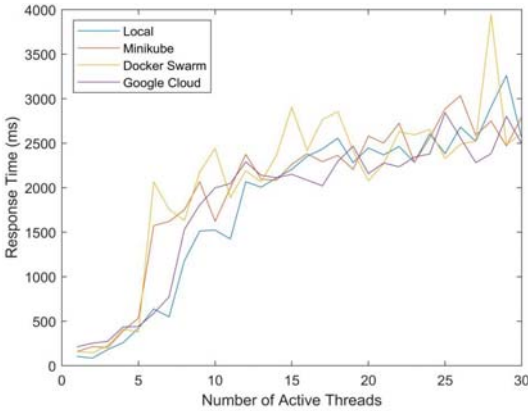


Fig. 2. Response time based on the number of active threads

The peak value and the saturation point differ from its type. The Google Cloud shows the lowest performance when the number of active threads are low. The reason for this is that the low number of active threads indicate that there is not much to process. In this case, network delay becomes the main issue, thus for Google Cloud being relatively far away, it is disadvantageous.

For this reason, in order to check the load balancing performance, it is worth to focus on the case when the number of active threads are high. In this case, the Google Cloud-based remote Kubernetes server and the Minikube-based local Kubernetes server shows higher transaction throughput than the Docker Swarm-based server.

Fig. 2 shows that the response time increases, when the number of users increases which indicates degradation of performance. After the number of users exceeds the specific point, the extent of growth decreases and seems to be saturated to the much higher response time as in the transaction throughput.

Among them, the Google Cloud shows the lowest response time in general, which means the highest performance. The Minikube follows the Google Cloud, and the Docker Swarm shows the lowest performance in terms of response time on average.

4. Conclusions

In this paper, the load balancing performance of Kubernetes and Docker Swarm was measured. As expected, Kubernetes outperformed Docker Swarm.

Acknowledgement

This research was supported by the Ministry of Science and ICT (MIST), South Korea, under the Information Technology Research Center (ITRC) support program (IITP-2020-2018-0-01799) supervised by the Institute for Information & communications Technology Planning & Evaluation (IITP).

References

[1] C. Pahl, A. Brogi, J. Soldani, and P. Jamshidi, “Cloud Container Technologies: A state-of-the-Art Review,” *IEEE Trans. Cloud Comput.* Vol. 7, No. 3, pp. 677-692, Jul.-Sep. 2019.

[2] E. Truyen, D. Landuyt, B. Lagaisse, and W. Joosen. “Performance overhead of container orchestration frameworks for management of multi-tenant database deployments.” *in Proc. the 34th ACM/SIGAPP Symp. Appl. Comput.*, Apr. 2019. pp. 156-159.

KSII The 12th International Conference on Internet (ICONI) 2020.

A Study on Cross-Domain Data Integration Strategy in Post-Pandemic Era

Jeonghoon Lee*
Division of Policy, Korea Institute of Science and Technology Information
Daejeon, South Korea
[e-mail: jhoon@kisti.re.kr]

Abstract

The increase of data and the development of information and communication technology motivate a data-driven technology paradigm shift, and at the center of this shift is open science, which is recognized as an innovation tool for solving global problems and sustaining social and economic development, and reflected as a national policy. Various problems such as climate change, disaster risk reduction, and sustainable development, which human society is facing in the 21st century, can be solved only through research collaboration between domains, and for this, cross-domain data sharing and utilization must be possible. At the time of the primary results of the global open science policies focusing the open data, the world is being forced to a radical digital transformation by a pandemic triggered by Covid-19. In this study, we propose a strategy for establishing a core interoperability and a method for integrating data between domains, which is the core of a data-intensive science ecosystem amid voluntary isolation of the global national society caused by the pandemic. These are reflected in the science and technology strategies of the CODATA and international open science groups, and are being promoted to build a data-driven digital infrastructure to respond to rapid social changes and global transformation.

***Keywords*:** Data Integration, Cross domain, Metadata, Science and technology policy and strategy, Covid-19, Post Pandemic, Global transformation

1. Introduction

We are living in an era where science and technology become the driving force of social change more than ever in human history. Data-driven technologies such as IoT, cloud computing, big data, mobile, and AI are at the base of the intelligence revolution. In particular, in line with the change to a data-driven society, the open science, focusing on open data, is drawing attention as a method for sustainable innovation growth of the nations. Efficient knowledge reproduction and acquisition became possible through ICT and accumulated digital research resources. Through such data-driven research collaboration, we believe that it is possible to find clues to solve the global social problems facing humanity such as climate change and disaster risk reduction, and to obtain the driving force for the sustainable development of each nation.

The pandemic triggered by Covid-19 is both a crisis and an extreme event that accelerates the transition to a new future society. It changed the environment more comprehensively and radically than any existing factors, and it became the driving force for changes in individual perception and behavior [1]. Problems facing the global world can be solved through collaborative research that seeks to understand complex systems through large-scale machine assisted analysis. However, the current capacity to combine heterogeneous data across and within domains is very limited. In particular, the cost incurred by the use of inconsistent data occupies most of research expenditures [2]. This

* The author is a member of CODATA International Data Policy Committee.

study examines the data issues in the post-pandemic era, and proposes a strategy for establishing core interoperability and a method for integrating data between domains, which is the core of a data ecosystem.

2. Data Issues in the Post-Pandemic Era

The biggest change caused by Covid-19 is that social distancing is expanding to the phenomenon of voluntary isolation of the national society and localization of information and technology beyond the level of preventing infectious diseases between individuals. In the face of global problems that require collaboration, polarization and uncertainty among domains, nations, continents are increasing. First, the conflict over information sovereignty in the data economy era is intensifying. In addition to the protection of data at the national level, self-determination of personal information is being considered in various fields. This is raising the need to establish a data distribution infrastructure that guarantees data localization and the decision-making right of the information subject. Second, it is expected that socioeconomic inequality will intensify due to the digital divide. Since the digital divide is changing from the divide in access of data to the divide in the level of information use [3], In addition, a curation system that preserves data and infrastructure as a single digital object should be established to ensure continuous accessibility of data.

3. A Cross-Domain Data Integration Infrastructure Design

A key enabler of data-driven research is to build an ecosystem that produces FAIR (Findable, Accessible, Interoperable and Reusable) data for machines as well as humans. For this, the data must be well-documented, have sufficient metadata, and processed so that human understandable meanings can be extracted from the results of complex phenomena and it can be used for future analysis and processing.

A cross-domain data integration infrastructure (CD2I2) is an interface for providing inter-operability between various existing data types and standards, and has the following basic features. Domain-independence, Atomic Data- Oriented Data Description, Data Provenance and Process, Foundational Metadata, Interoperability and Alignment with other Standards. First, CD2I2, not attempt to be a domain model of semantics, is based on the general structure and aspects of the object being described, rather than a model limited to the life cycle of a particular domain. Second, data description is based on atomic datum. They can share data transfer methods between domains by assigning roles to datum according to the structure of the data set: wide data, long data, dimensional data, key-value data, streaming data, big data, registers, instrument data, etc. Third, in order to fully understand the data, it is necessary to know the processing and conversion process of the data. To accurately understand how individual data is represented and processed in different data sets, it is necessary to explain the data provenance. At last, terminology for the same component across domains is established and it must be interoperated with other standards and models in order to implement across a wide variety of technology platforms.

4. Concluding Remarks

The global digital transformation is rapidly accelerating, even under the non-face-to-face and distancing across society caused by Covid-19, and the need for collaboration across domains is increasing. In order to interoperate data amid the localization of information and digital divide, data integration between domains is necessary, and efforts are being made internationally to establish this in the form of a digital infrastructure. In accordance with the a cross-domain data integration infrastructure design examined in this study, the establishment of collaboration and application cases are currently in progress, and various policy and technical studies for mobilizing domains and breaking down silos will be conducted with international partners.

References

[1] J. Yoon, “After Covid-19, changes in the new normal and future society,” *STEPI Future Horizon*, Vol 46, 2020.

[2] S&T Committee, “Balance and effectiveness of research and innovation spending,” *Report for House of Commons of the UK*, 2019.

[3] NIA, “The report on the digital divide,” *Report for Ministry of Science and ICT*, 2019.

Expansion of English Corpus for Training by Sentence Transformation Rules

[1]Sungho Shin and [1,2]Sa-Kwang Song
[1]Research Data Sharing Center, Korea Institute of Science and Technology Information
[2]Department of Big Data Analysis, Korea University of Science and Technology
Daejeon, South Korea
[e-mail: {maximus74, esmallj}@kisti.re.kr]

Abstract

Machine learning performs well for information extraction tasks, and the training corpus plays an important role with regard to the accuracy of information extraction. Because considerable costs are incurred when constructing a training corpus that is sufficiently large for studies and practical applications, recent studies have been focusing on statistical approaches which perform well with minimum training corpus. However, it is difficult to guarantee higher accurate results with statistical ways than using fully manually constructed training corpora. In this study, a small initial training corpus is expanded by rules, such that the training accuracy reaches to 89% of that of the entire corpus.

***Keywords*:** Information extraction, training corpus, sentence transformation, corpus expansion

1. Introduction

This study aims at reducing the effort and time required for building a training corpus. It differs from conventional studies that depend on initial small training corpus as the training corpus. Our proposed method increases the size of the training corpus by transforming the sentences from the initial training corpus. The main idea behind the sentence transformation is a structural transformation and lexicon expansion of sentences because the features used in machine learning are mainly divided into lexical and syntactical features. When changes are made to the structure and the lexicon of sentences in an initial training corpus, learning models will recognize the transformed sentences as sentences that are different from the sentence in the initial training corpus.

To verify the method proposed in this study, two common information-extraction tasks were applied: named-entity recognition and relation extraction. The former was limited to technology names, and the latter was restricted to three semantic relations.

2. Research Design

2.1 Model

Since a learning model meets various sentences in structure and lexicon during prediction, it is likely that a learning model with a training corpus composed of sentences containing various structures and lexicons will perform well.

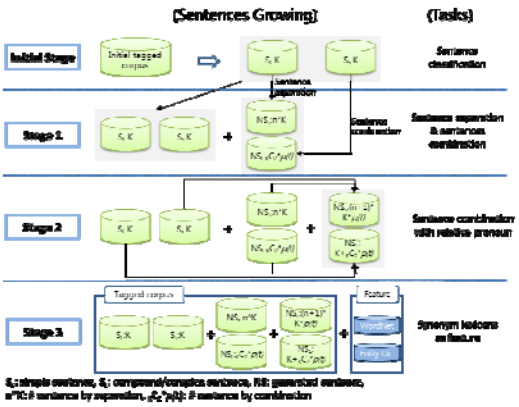


Fig. 1. Proposed model for corpus expansion

Fig. 1 shows the corpus expansion model proposed in this study. This model focuses on diversifying the sentence structure and the lexicon of sentences used for training in order to obtain the similar result with using the entire corpus.

The corpus expansion model consists of four stages. The initial stage is a preparation stage for expanding the corpus. Each sentence from the initial training corpus is classified as a simple sentence, compound sentence, or complex sentence. In Stage 1, many new simple sentences are generated by separating compound sentences and complex sentences respectively. Conversely, a simple sentence is combined together with another simple sentence to make a new compound sentence or a new complex sentence, which is the process of sentence combination.

In Stage 2, the sentences are connected with each other using relative pronouns (RPs). Each sentence becomes a main clause, and all of the simple sentences apart from this sentence become candidates for a RP clause. The criterion for becoming a RP clause and connecting to a main clause is having a tagged technology name as a subject in the main clause. After increasing the number of sentences in the training corpus, the lexicon is expanded by exploiting a synonym dictionary for critical words.

2.2 Sentences Verification

Sentences generated from initial training corpus through separation or combination are unverified sentences. For verification, the PCFG method was used, which is used often in conventional studies. In the PCFG method, which word will appear after each word is calculated in advance as a probability value by analyzing the existing sentences. A parsing model of the Stanford parser calculates the PCFG value when a sentence is given as an input value.

To know the PCFG values of general sentences, about 100,000 sentences were randomly collected from Wikipedia, and the probability distribution was estimated using the PCFG scores of the sentences. The confidence interval was 95%. When the newly generated sentences belonged to this distribution, the verification was considered to have been completed.

3. Experiment and Results

The experiment was performed from two aspects: first, the performance of the expanded training corpus, compared to the training corpus constructed by entirely manual work; and second, the size of the initial training corpus at the start. For the two experiments, 5 kinds of training corpora were prepared; they were initial 20%, initial 40%, initial 100% training corpus, initial 20% with transformation, and initial 40% with transformation.

For the accurate performance evaluation, the ratio of three relations, namely compete, similar, and elementOf, in each training corpus was identically set to 3:2:5. As tools for the performance evaluation of the relation extraction, precision, recall, and F1 score were used. Structural Support Vector Machine (SSVM) was applied for the information extraction task. This model is known for its strength in structural analysis such as sequence labeling and syntactic analysis [1]. **Table 1** shows the results of the performance evaluation.

Table 1. Experiment result

Corpus	# Sentence	Precision	Recall	F1 score
100% initial	7530	82.273	70.257	75.791
40% initial	3020	68.394	64.706	66.499
40% transformed	4861	70.732	65.022	67.757
20% initial	1513	40.909	29.717	34.426
20% transformed	2287	15.467	27.358	19.761

4. Conclusions

In this study, a training corpus expansion model is proposed and it shows two implications through experiment. First, the size of initial training corpus which is to be transformed for expansion must not be small. 40% of the sufficient corpus is at least recommended. Second, the proposed model makes improvement of accuracy in information extraction and reaches to 89% of the accuracy of the entire corpus which is all done by human. It could contribute to reduce the cost of building a training corpus.

References

[1] C. Lee et al., “Balanced Korean word spacing with structural SVM,” *In EMNLP*, pp. 875-879, 2014.

A Preliminary Study on UI/UX Design Evaluation Guidelines for XR Contents

Eeksu Leem[1], Jeongwook, Seo[2], Mingoo, Kang[2*], Hyoseok Yoon[3], Sangyep Nam[4], and Hyunsik Kim[5]

[1]MR Media Artech Center, Hongik University, Seoul South Korea
[2]Dept. of IT Contents, Hanshin University, Gyeonggi South Korea
[3]School of Computer Engineering, Hanshin University, Gyeonggi South Korea.
[4]Dept. of Electronic Information and Communication Eng., Kookje University, Gyeonggi South Korea.
[5]Contents Convergence Research Center, KETI, Seoul, South Korea
[e-mail : remixleem@hongik.ac.kr, jwseo@hs.ac.kr, kangmg@hs.ac.kr, hsyoon@hs.ac.kr, synam58@gmail.com, hskim@keti.re.kr]
*Corresponding author: Min-Goo Kang

Abstract

In this paper, we presented basic evaluation guidelines on UI/UX (User Interface and User eXperience) for XR (eXtended Reality) stream image contents. In addition, the guideline of UX/UI design and usability of XR contents were with the quality control of digital twin-based XR contents developed for remote collaboration, process performance, and maintenance purposes in the manufacturing, as well as an evaluation system that can effectively perform it.

***Keywords*:** UX/UI design, eXtended Reality, digital twin, quality control

1. Introduction

Evaluation of the UI/UX (User Interface and User eXperience) design for virtual reality (VR) contents has comprehensively measured these factors such as (1) usability, (2) usefulness, (3) satisfaction.

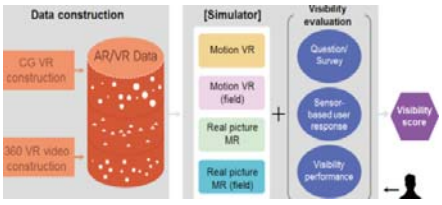


Fig. 1. XR Content of UI/UX Design and Evaluation

However, UI/UX design evaluation for virtual XR(VR+AR+MR) contents should be extended evaluation area, such as safety usage, user-health because XR contents in digital twin systems interact with physical reality [1].

In this paper, we studied UI/UX design evaluation guidelines based on Aaron Walter's user demand layer and suggest a web-based evaluation system for XR content.

In terms of evaluation, evaluation guidelines will be expected to use of quality control for enhanced satisfaction of further digital twin-based XR applications [2].

2. UI/UX Design and Web-based Evaluation System for XR Contents

2.1 Evaluation of UI/UX design(XR contents)

The reliability of digital twin-based XR content must be ensured through certain quality assurance efforts in the manufacturing area. In

This research was supported by the MSIT Korea, under XR-Convergence project supervised by the NIPA [V0514-20-1011, Digital operation system for XR-based manufacturing facilities].

particular, the guideline of UX/UI design and usability is the main factors that determine the quality of XR contents[3].

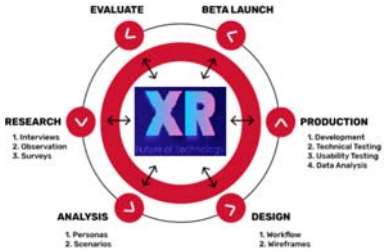


Fig. 2. Evaluation of UI/UX design(XR contents)

2.2 Data collection, guideline for evaluation based on international standards

It is necessary to utilize international standards recognized for objectivity and quality for a large number of user groups in order to carry out objective evaluations and obtain reliability and validity of measurement results

Because the UX/UI design and usability of content is greatly influenced by the user's subjectivity and sensibility.

Fig. 3. Process of XR contents and UI/UX design

2.3 Necessity of guidelines for evaluating the usability of XR contents

'XR_use and production safety guidelines' are used to evaluate the usability of XR contents, but only the satisfaction of XR contents is measured as the scope of application is targeted at devices for games and amusement facilities.

Because the purpose of XR content in the manufacturing sector is to perform the actual task, it is essential to measure the suitability of the purpose, such as whether the use of XR content interferes with the work due to cognitive overload. In addition, the environment in which manufacturing XR content is operated is often a real work environment, not a risk factor. The importance of safety factors is very high compared to the XR content for amusement, too.

2.4 Web-based XR content usability system

As the fandom situation such as Corona 19 continues, it is expected that it will be difficult to carry out the actual research if the existing face-to-face usability evaluation system is established. Therefore, it is necessary to create a foundation for both face-to-face and untacted evaluation system depending on the situation[3].

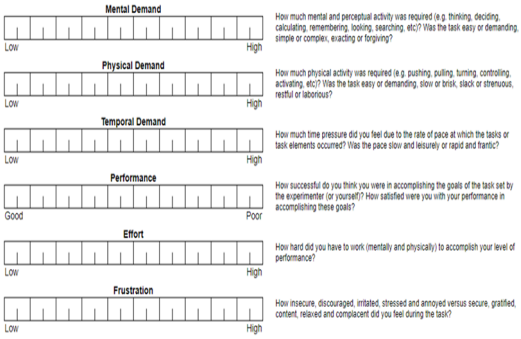


Fig. 4. Web-based assessment items and evaluation factors for XR content usability system

3. Conclusion

In this paper, a scripted performance protocol was developed and presented for the facilitation of the progress of usability assessment. As a result, a diagram of the role-specific, step-by-step performance scenarios was specified for collection elements, and tools, as well as the behavior of the experimenters

References

[1] Y.K., Jung, "Development of VR Fire-extinguishing Experience Education Contents Using UX Design Methodology," *International Journal of Contents*, Vol.17, No.3 pp.222-230, 2017.

[2] J.W. Park et al., "Design and implementation of VR-based life care contents for attention deficit hyperactivity disorder (ADHD)," *Korean Institute of Next Generation Computing*, Vol.15, No.5 pp.84-91, 2019.

[3] Y. J. Oh, E. B. Kang, and J. H. Suh, "Improvement Suggestions for Mobile News Report Application UI", Proceedings of HCI KOREA 2017, Gangwon, p637-640, (2017)

Emotion analysis and behavior prediction using layered voice analysis

Minjae Park[1], Sanghyun Paik[2] and Kwanghoon Pio Kim[3,4*]
[1]Division of Computer Information, Daelim University College
Gyeonggi, South Korea
[2]Business Planning Division, MIRAESON Co., Ltd.
Seoul, South Korea
[3]Division of Computer Science and Engineering, Kyonggi University
Gyeonggi, South Korea
[4]Contents Convergence Software Research Center, Kyonggi University
Gyeonggi, South Korea
[e-mail:, mjpark@daelim.ac.kr, billy@miraeson.com, kwang@kyonggi.ac.kr]
*Corresponding author: Kwanghoon Pio Kim

Abstract

Solving mental and emotional problems is emerging as a very important issue in modern society. We try to analyze and use voice data to solve mental and emotional problems. People's emotions are expressed in voice, and voice data can be accessed somewhat easily through conversations or phone calls. We would like to propose a method for analyzing emotions and predicting behavior by utilizing this voice data and using hierarchical voice analysis technology.

***Keywords*:** LVA, Layered Voice Analysis, Emotion Analysis

1. Introduction

In many cases, modern people live with emotional problems. Mental illness, depression caused by stress, and socioeconomic problems caused by suicide are known to be serious. If you look at the statistics on the mental health problems of Republic of Korean youth in their 20s, it can be seen that it is at a very serious level. In other words, mental health problems can become very important socially in the near future. In order to find these mental problems, we try to analyze data through voice information. Voice data is easy to access and is likely to contain various emotional information.

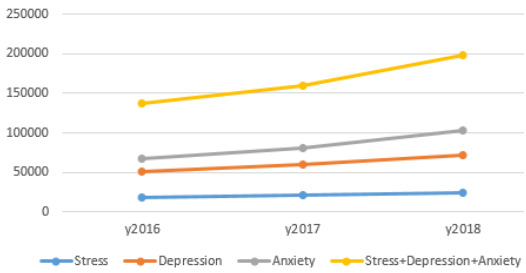


Fig. 1. Year 2016~2018 medical treatment status (20s)**,** Health Insurance Review & Assessment Service

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (grant number 2020R1A6A1A03040583).

2. Layered voice data analysis

We tried to examine the validity of the analyzed data by installing a system for sentiment analysis and analyzing voice data through calls.

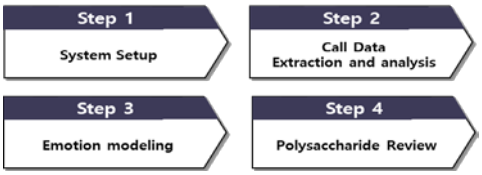


Fig. 2. Emotion analysis and behavior prediction process through voice data analysis

Based on the voice-based data, to obtain emotional numerical information and analyze the emotional information, a 4-step procedure is performed as follows.

1) Collect the speaker's voice to be analyzed. 2) Segment the sound source into segments (less than 2 seconds). 3) The fine pattern of the voice fragment is compared with the voice identifications possessed by the engine. 4) Classify emotion items based on the compared values.

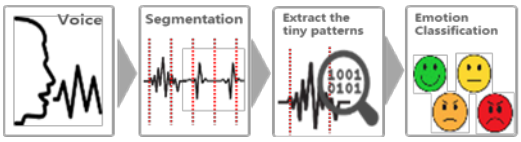


Fig. 3. Voice data analysis procedure

3. Experimental results

We tried to analyze the data based on data such as pleasure, energy, stress, anxiety, concentration, confidence, embarrassment, and anger.

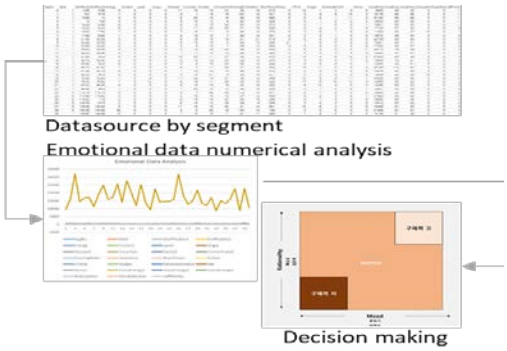


Fig. 3. Voice data analysis and behavior predition

4. Related works

Recently, computing technology for recognizing human emotions in various ways has been studied. The method[2] of comprehensively analyzing the pitch, intonation, and vocabulary of the voice is distinguished from us because we are based on the voice itself. In addition, research on human behavior[3], bio-signals[4], and emotional analysis through texts[5] written by users are also in progress. We mainly dealt with the recognition of voice itself, which is easy to access and has a wide range of applications.

5. Conclusions

In this paper, we attempted to recognize emotions through voice data and predict behavior through the recognized information. Voice data will contain various emotions, and we used a system known to be able to extract various emotions, quantified and extracted major emotion data, and tried to obtain meaningful results through the extracted information.

References

[1] 2016~2018 medical treatment status, http://www.ssunews.net/news/articleView.html?idxno=7529

[2] Hupp, Oliver, et al. "Beyond verbal scales: Measurement of emotions in advertising effectiveness research." Yearbook of Marketing and Consumer Research 6.4 (2008): 72-99.

[3] Huang, Weidong, et al. "Sens-e-Motion: Capturing and Visualising Emotional Status of Computer Users in Real Time." 2019 23rd International Conference in Information Visualization–Part II. IEEE, 2019.

[4] Hosseini, Seyyed Abed, and Mohammad Bagher Naghibi-Sistani. "Emotion recognition method using entropy analysis of EEG signals." International Journal of Image, Graphics and Signal Processing 3.5 (2011): 30.

[5] Al Marouf, Ahmed, et al. "Recognizing Language and Emotional Tone from Music Lyrics using IBM Watson Tone Analyzer." 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT). IEEE, 2019.

Method of Calibrating Inertial Measurement Unit for Recording Human Pose Data

Meejin Kim and Byung-Ha Park
Korea Electronics Technology Institute
Seoul, South Korea
[e-mail: mj_kim1023@keti.re.kr, bhpark@keti.re.kr]
*Corresponding author: Byung-Ha Park

Abstract

Since the Inertial Measurement Units (IMUs) provide the raw data based on each of the sensor-local coordinate systems, the researchers have necessary to align the data with the body-centered coordinate system for reconstructing the human pose. This paper proposes a method to calibrate 6 IMUs worn on the human body for recording human pose datasets. Our calibration method utilizes 3-axis accelerometer data measured from IMUs with MEMS sensors in two poses. Through the dominant axis in the accelerometer, we identify the direction in which gravity acts on the sensors and calculate the orientation from the global coordinate system of each sensor. We further demonstrate the visualization of the coordinate frames which shows the calibrated sensor data for the human model.

***Keywords*:** Calibration, Inertial Measurement Unit (IMU), Micro Electro-Mechanical Systems (MEMS)

1. Introduction

The demand for real-time 3D human pose estimation continues in a wide range of areas, including Virtual and Augmented Reality (VR/AR). However, the common marker-based motion capture systems require expensive equipment and large space that are difficult to use individual. For this reason, many researchers have attempted to simplify the method using inertial measurement units (IMUs). Existing commercial systems typically require 17 sensors for motion capture [1]. Several researchers only used 6 IMUs to compute full-body 3D poses [2, 3] and fused video input and IMUs data [4, 5]. Although the difference of the input data, these studies commonly involved the process of recording pose data using IMUs.

To record the pose data, a calibration process is required to align the coordinate system of each IMUs to the body. In this paper, we present a convenient method of calibration using 3-axis acceleration data in two predefined poses, A-pose and T-pose. Then, we calculate the orientation from the global coordinate system of each sensor through the direction of gravity and collected values. We expect it will contribute to the convenience of pose research.

2. Predetermined Pose based Calibration Method

2.1 Sensor Alignment

We use 6 LPMS-B2 IMU sensors containing 3-axis accelerometers, gyroscopes, and

This work was supported by Korea Evaluation Institute of Industrial Technology(KEIT) funded by the Korea government(MOTIE). [Project No:20009799 / Name: Development and verification of service innovation technology for small and mediumsized retail business]

magnetometers. The sensors combine the data from three sensing units using the Kalman filter to calculate the orientation and provide raw data based on sensor-local coordinates.

For our calibration method, we use only the data from the right-hand sensor, called *basis sensor*. Hence, we arrange all the sensors in the same direction on the flatten place to obtain rotation to the sensor orientation from the basis sensor orientation R^{BS}.

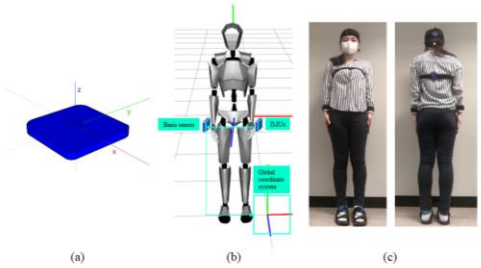


Fig. 1. (a) Sensor-local coordinate system. (b) calibration overview. (c) Sensor placement during calibration.

2.2 Calibration from Predetermined Pose

We measure the 3-axis acceleration data in two determined poses, putting the arms down straightly (A-pose) and spreading the arms out horizontally (T-pose). Assuming the root orientation of the body as the global coordinate system, by cause of that gravity does not always change, the most dominant axis of acceleration data of basis sensor should be negative y-axis $(-y)$ in A-pose and negative x-axis $(-x)$ in T-pose.

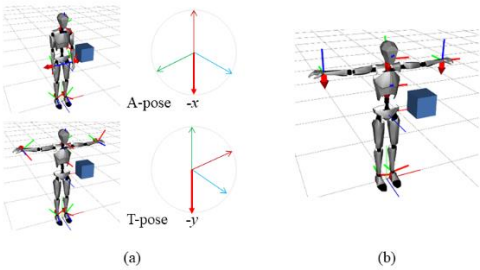


Fig. 2. (a) R^{GB} can be calculated through the direction of gravity applied to the basis sensor. (b) The sensor orientation in T-pose with our calibration method.

With this assumption, we can get a rotation R^{GB} that causes the raw data of the basis sensor to follow the body-centric coordinates systems. In the following, we will assume R^{GS} which indicates all IMUs coordinates in the global coordinate system using the transformation rule when the body model is defined in the global coordinate system.

$$R_t^{GS} = R^{GB} R_t^{BS} \tag{1}$$

3. Conclusions

We presented an accessible method to align IMUs sensor-local coordinates to the global coordinate system using two defined poses and gravity. We would use it to record pose data for estimating 3D human body pose in real-time, used with interactive technologies.

References

[1] Roetenberg, D., Luinge, H., & Slycke, P. (2007). Moven: Full 6dof human motion tracking using miniature inertial sensors. *Xsen Technologies, December*, *2*(3), 8.

[2] von Marcard, T., Rosenhahn, B., Black, M. J., & Pons- Moll, G. (2017, May). Sparse inertial poser: Automatic 3d human pose estimation from sparse imus. In *Computer Graphics Forum* (Vol. 36, No. 2, pp. 349-360).

[3] Huang, Y., Kaufmann, M., Aksan, E., Black, M. J., Hilliges, O., & Pons-Moll, G. (2018). Deep inertial poser: learning to reconstruct human pose from sparse inertial measurements in real time. *ACM Transactions on Graphics (TOG)*, *37*(6), 1-15.

[4] Von Marcard, T., Pons-Moll, G., & Rosenhahn, B. (2016). Human pose estimation from video and imus. *IEEE transactions on pattern analysis and machine intelligence*, *38*(8), 1533-1547.

[5] Trumble, M., Gilbert, A., Malleson, C., Hilton, A., & Collomosse, J. P. (2017, September). Total Capture: 3D Human Pose Estimation Fusing Video and Inertial Sensors. In *BMVC* (Vol. 2, No. 5, pp. 1-13).

KSII The 12th International Conference on Internet (ICONI) 2020.

Two Hands Segmentation from Depth Images with a Neural Network

Sukwon Lee, Sungoh Jung and Byung-Ha Park
VR/AR Research Center, Korea Electronics Technology Institute
Seoul, South Korea
[e-mail: sukwonlee@keti.re.kr, pkcsoh88@keti.re.kr, bhpark@keti.re.kr]
*Corresponding author: Sukwon Lee

Abstract

We propose a method that automatically segments two hands from depth images with the neural network. The structure of the proposed neural network is similar to an auto-encoder, which stacks several deconvolution layers after convolution layers, but the output is pixel-based labels. For the training network, we build an automatically generated dataset with colored gloves and background. By aligning viewports between the RGB sensor and depth sensor, colored gloves can provide the information whether each depth pixel belongs to the hands or background. We experiment with the performance of our method from naïve situations to highly occluded one by itself.

***Keywords*:** depth segmentation, two-hand segmentation, convolutional neural network

1. Introduction

The interaction with a natural interface such as hands are enlightened its importance by making controlling the application easy. As the Head Mounted Display (HMD) is immersed broadly, the hand is the rightful interface but the recognition of the user's intention is hard since its complexity and ambiguity. For interacting with the virtual object with hands, the estimation of the hand pose is essential. In addition, the pose of both hands will be flourishing the interaction. Several works [2][3] researched on two hand detection and pose estimation with RGB-D cameras, which directly detect two hand poses. On the other hand, works [1] related to hand pose estimation are focused on the improvement of only the one hand pose. Therefore, we propose the hand segmentation method from the depth image, which is able to apply one hand pose estimation after the segmentation. To this end, we implemented the Convolution Neural Network (CNN) that forms as the auto-encoder, and the output is pixel-based segmentation labels. For the training, we gathered the data that is labeled by the color, which can be reduced the labor for labeling. After training, our method automatically segments both hands from a single depth image.

2. Segmentation

2.1 Segmentation with colors

Our method takes a depth image of what each pixel represents the distance from the camera as the input, and a labeled data as the segmented output. For the training, this method needs depth images and corresponding labeled data. However, making the labeled data is the most laboring process, which makes it impossible to build training data because of the size of the dataset that is over thousands in our case. To deal with this problem, we segment between each semantics by colors; For example, blue colored

This work was supported by Korea Evaluation Institute of Industrial Technology(KEIT) funded by the Korea government(MOTIE). [Project No:20004356 / Name: Development of Industrial Safety Virtual Training System to Cope With Complex Risk Situation in Industrial Field]

pixels represent the background, the red one is the left hand, the white one is the right hand.

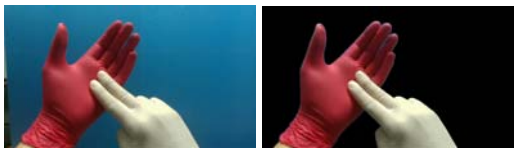


Fig. 1. (Left) a color image that segments by the color. (Right) The result after deleting the background.

For easy splitting, the colored gloves and the background have higher hue values. With these images, we first detect the pixel candidates that are included in each semantic area. After that, Graphcut algorithm is used for giving the initial guess of the labeled pixel. In fact, these initial guesses are not accurate to give it directly to the training. Thus, we implemented an interactive tool to refine label pixels by giving a rough correction. With little corrections, we re-run the Graphcut algorithm, and then results reach higher accuracy.

2.2 Auto-encoder

We implement the convolution neural network, which formulate as an auto-encoder by stacking convolution and de-convolution layers serially. The encoder part of our network consists of 3 convolution layers, on which the size is reduced as it goes deeper. After the encoder part, the decoder part is followed, on which the size of each layer increases. The input layer takes the depth image and the output layer gives the labeled data.

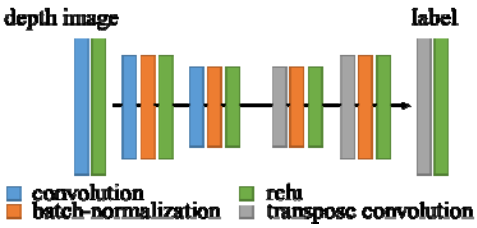


Fig. 2. Auto-encoder implemented for out method.

3. Results

We measure the training loss for the training dataset and validation dataset.

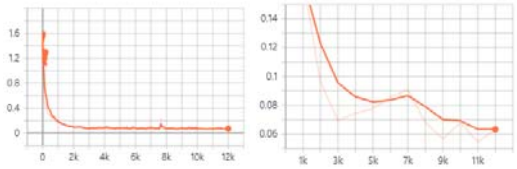


Fig. 3. (Left) training loss, (Right) validation loss As figures are shown, the loss converges after 2000 iteration.

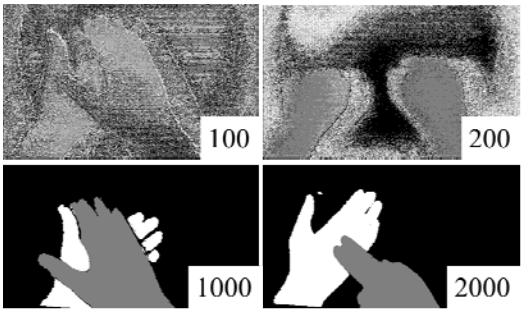


Fig. 4. Segmentation results. The right bottom number represents the training iterations.

4. Conclusions

In this paper, we present a two-hand segmentation method from a single depth data with CNN.

References

[1] Tkach, A., Tagliasacchi, A., Remelli, E., Pauly, M., & Fitzgibbon, A, "Online generative model personalization for hand tracking," *ACM Transactions on Graphics (ToG)*, *36*(6), 1-11, 2017.

[2] Mueller, F. et al., "Real-time pose and shape reconstruction of two interacting hands with a single depth camera," *ACM Transactions on Graphics (TOG)*, *38*(4), 1-1, 2019.

[3] Bojja, A. K et al., "Handseg: An automatically labeled dataset for hand segmentation from depth images," *16th Conference on Computer and Robot Vision (CRV)* (pp. 151-158). IEEE., 2019.

KSII The 12th International Conference on Internet (ICONI) 2020.

Remote Synchronization for Identifying the LED Markers

Sukwon Lee, Byung-Ha Park
VR/AR Research Center, Korea Electronics Technology Institute
Seoul, South Korea
[e-mail: sukwonlee@keti.re.kr, bhpark34@keti.re.kr]
*Corresponding author: Sukwon Lee

Abstract

We propose a synchronization method between systems remotely separated by simple infrared LED and receiver, and we use this system for identifying the LED unique ID. We use the infrared receiver that is commonly used for sending a command to home appliances. As a marker, the LED firstly transmits a synchronization code with a frequency pulsed about 38kHz, and the identification number sending part follows. The identification of LED is 8-bit data, and sending it by varying the duration of the 'on' part of the LED to distinguish the brightness of the marker. As a receiver, we use an infrared imager and infrared receiver that sensing the signal in which the frequency is around 38kHz. The infrared imager takes the image after the receiver got the synchronization code from the marker. Because there is a latency between the computer and the receiver, the LED ID can be computed after the calibration sequence. We build and test this system using Arduino and RealSense™ camera.

***Keywords*:** Active marker, remote synchronization, infrared receiver

1. Introduction

Many industries use the motion capture data for making animation, analyzing human movement, and interacting with the virtual space. Despite its importance, the motion capture process is hard to access because of the cost of the capture system. To get accurate motion data, the system must correctly identify the markers in the case of the outside-in capture. Passive markers are cheap and commonly used for identifying marker but there is an ambiguity on the correspondence at each frame. To end this, an active marker[1][2][3] is introduced, which is a self-emissive identification to the capture system. Thus, accuracy is improved by removing the ambiguity of the correspondence. We introduce an active marker-based motion capture system that communicates in a remote manner between marker and system. For making the synchronization on the air, we use an infrared LED and receiver that is used for controlling home appliances, and also very cheap. We build the system with Arduino and a RealSense™ camera.

2. Remote synchronization

2.1 Communicate protocol

We define a protocol to communicate with the system. Our protocol consists of two parts. The first part is the synchronization signal followed by the identification sending part. The synchronization signal can be detected by the infrared receiver by flickering at a high

This work was supported by Korea Evaluation Institute of Industrial Technology(KEIT) funded by the Korea government(MOTIE). [Project No:20008948 / Name: Development of virtual training service system for cooperating flight crew team response to abnormal flight conditions]

frequency of about 33~58 kHz. This signal will show a continuous light from the camera.

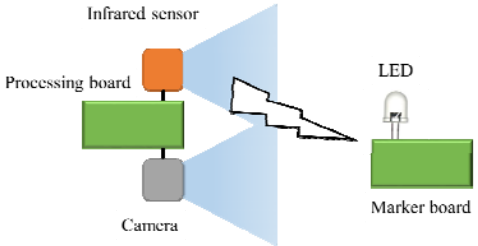


Fig. 1. System overview.

The identification of the marker is composed of 10-bits and sending serially matched with the synchronization signal. Unlike the synchronization signal, the identification signal is interpreted with the camera by comparing the brightness between incident frames. Thus, the latency to get the identification is determined by the specification of the camera. In our experiments, we use D435 camera, which takes 90 frames per second, and then latency is about 0.1 seconds. However, this value is negligible because latency only exists when the tracking is on starting. We apply a Kalman filter to reduce mistracking.

2.2 Decoding

The synchronization signal represents the timing of the signal and its mode of the signal. As soon as an infrared signal is received, the processing board signals to the camera to open the shutter. After 1ms, the shutter will be closed and sending a captured infrared image to the processing board. Then, the board identifies what ID of each marker is shown on the image by comparing previous frames. When we look at several images in serial, a marker is shown as the light in a circle shape and it varies the size along with frames according to the ID bit value. Because the marker sends a bit value by LED brightness, we can decode the bit after 10 frames are received in a row.

3. Implementation

We use a personal computer (PC) as a processing board and Arduino board as a marker board. We choose RealSense™ D435 as the camera because it can take a synchronization signal from the outside. To generate a signal, we build an additional board on the camera, which has an infrared receiver on it. In front of the camera, we glue the band-pass filter that passes 850nm.

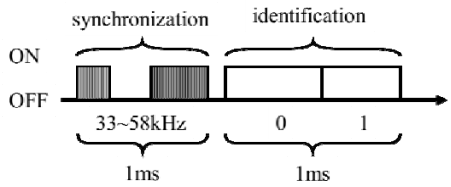


Fig. 2. Communication protocol.

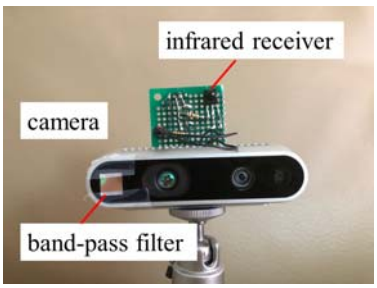


Fig. 3. The camera and the additional board for signaling to the camera.

4. Conclusions

In this paper, we introduce a remote synchronization system for identifying marker ID even if there is no connection. We build our system with affordable components, which can help reduce the cost to make a capture system.

References

[1] Gorse, L., Löffler, C., Mutschler, C., & Philippsen, M., "Optical Camera Communication for Active Marker Identification in Camera-based Positioning Systems," *15th Workshop on Positioning, Navigation and Communications (WPNC)*, (pp. 1-6) 2018.

[2] Urtans, E., & Nikitenko, A., "Active infrared markers for augmented and virtual reality," *ENGINEERING FOR RURAL DEVELOPMENT*, 2016.

[3] Schmid, S., Corbellini, G., Mangold, S., & Gross, T. R., "Continuous synchronization for LED-to-LED visible light communication networks," In *2014 3rd International Workshop in Optical Wireless Communications (IWOW)* (pp. 45-49). IEEE, 2014.

KSII The 12th International Conference on Internet (ICONI) 2020.

The study of complex sensor using the IoT sensor and artificial intelligence model

Youngmin Ji and Dongwoo Kwon
Energy IT Convergence Research Center, Korea Electronics Technology Institute (KETI)
Seongnam, South Korea
[e-mail: ym.ji@keti.re.kr, dwkwon@keti.re.kr]
*Corresponding author: Dongwoo Kwon

Abstract

IoT, big data, and artificial intelligence technologies are advancing day by day, but until now, there is still a lack of convenient ways to converge technologies in each field. It is possible to infer new meanings of artificial intelligence learning models by analyzing IoT sensor data based on past historical data, but it is not easy to combine individual sensor data with new artificial intelligence models. In this paper, we introduce a method for developing a new service by developing an artificial intelligence model for indoor occupancy detection based on the data acquired from an IoT sensor and uploading a trained model on the cloud so that various sensors can easily compose artificial intelligence models

***Keywords*:** Artificial Intelligence, Bigdata, Complex Sensor, Internet of Things

1. Introduction

Recently, the popularity of IoT, big data, and artificial intelligence technologies have not yet cooled down. However, it is a reality that those who want to use it in the actual field do not know this technology well, and how to apply it to the field is very difficult. Therefore, in this paper, we introduce a method of composing an artificial intelligence and creating an IoT service using an artificial intelligence model through a drag-and-drop user interface in a cloud-based environment.

2. Internet of Things with Artificial Intelligence models

2.1 IoT Environment Sensor

IoT Environment sensor[1] developed for monitoring the indoor environment and was equipped with seven sensors for temperature, humidity, CO2, illuminance, sound, organic compound, and motion. The sensing function of this device is tested for interoperability by using WoT standard[2] technology.

2.2 Artificial Neural Network Model

Various data received from the environmental sensor can act as a necessary factor for indoor occupancy detection. However, it is difficult for a person to judge these factors individually, so a model is constructed through an artificial intelligence neural network.

We provided a deep learning-based camera as the correct answer to teach the occupancy detection neural network model and learned by the sensing data of the IoT environment sensor. The training data that we used were the deep learning recognition results recognized every 5 minutes and the average of the 5-minute sensing values measured by the IoT environment sensor.

In this paper, JavaScript-based brain.js[3] is used as an artificial intelligence neural network framework, and the trained model is stored in the

This work was supported by the Korea Sports Promotion Foundation (KSP0) and the Ministry of Culture, Sports and Tourism (MCST) of the Republic of Korea (No. S2020-04-02).

form as shown in **Fig. 1**. This model consists of 9 inputs, 10x10x2 hidden layers, and 1 output, and the output is 0~1, indicating the probability of occupancy.

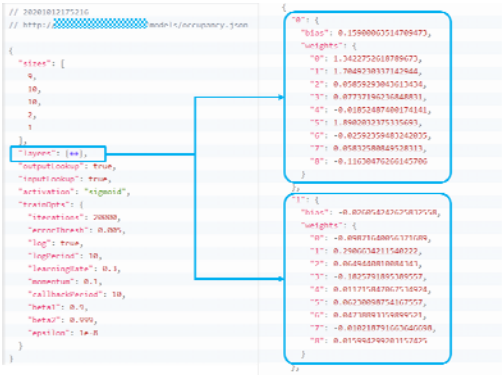


Fig. 1. Trained Neural Network Model

3. IoT Service Composition with AI model

We used Node-RED[4] as the basic tool for service composition. In **Fig. 2**, we developed nodes that support Things Description on this tool, and easily connected devices based on W3C(World Wide Web) Web of Things standard with drag and drop.

Fig. 2. IoT connection based on the web of things

The node reads the IoT sensing value every 5 seconds, converts it to JSON, and translates it into a neural network input form using a function node in **Fig. 4**.

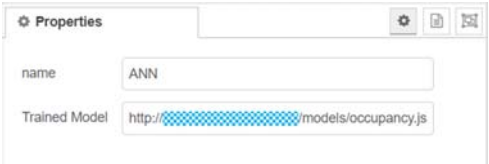


Fig. 3. Trained Neural Network Model

Fig. 3 shows the nodes that can pull trained artificial intelligence neural network model from the cloud. We compose this model with IoT sensing data to create the service based on drag and drop with Node-RED flow.

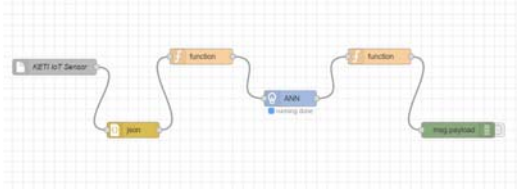


Fig. 4. IoT Service Composition Flow in Node-RED.

Fig. 4 shows the Node-RED flow that finally composes the service. This flow is a service that takes IoT sensing data every 5 seconds and infers the probability of indoor occupancy through a neural network. In this way, it is possible to easily apply artificial intelligence services by linking data from IoT sensors.

4. Conclusions

In this paper, we show that users can easily configure artificial intelligence services with IoT using Node-RED tool even if they do not know about artificial intelligence neural networks. In future research, we plan to study the concept of service composition that can be updated periodically and automatically reflected in the service when the model is updated.

References

[1] Y. Ji, K. Ok and D. Kwon, "The indoor environment monitoring system for intelligent buildings using Wifi mesh based Internet of things", *Proc. 2019 17th Annu. Int. Conf. Mobile Systems Applications and Services (MobiSys)*, pp. 689-690, Jun. 2019.

[2] S Kaebisch, T Kamiya, M McColl, V Charpenay, M Kovatsch, "W3C(World Wide Web) Web of Things (WoT) Thing Description", *W3C Recommendation 9 April 2020*. https://www.w3.org/TR/wot-thing-description/

[3] Brain.js:"GPU accelerated Neural networks in JavaScript", https://brain.js.org/

[4] Node-RED, "Low-code programming for event-driven applications", http://nodered.org

Sensor Configuration-based IoT Environmental Sensor Device Management

Dongwoo Kwon[1], **Kisu Ok**[2], **and Youngmin Ji**[3]
Energy IT Convergence Research Center, Korea Electronics Technology Institute (KETI) [1,3]
Seongnam, South Korea
Dept. of Electrical and Computer Engineering, Seoul National University[2]
Seoul, South Korea
[e-mail: {dwkwon, ymji}@keti.re.kr[1,3], ksok@redwood.snu.ac.kr[2]]
*Corresponding author: Youngmin Ji

Abstract

Recently, the deployment of Internet of things (IoT) environmental sensors in various types of buildings and factories is increasing in order to reduce energy use and improve energy management efficiency. The configuration information of these sensor devices varies depending on the installation location, purpose, and surrounding environment of the device. Therefore, it is difficult to distribute sensor devices to multiple sites and to manage software images of already installed devices. To solve the problem, this paper proposes an IoT environmental sensor device management scheme based on separate sensor configuration in a software configuration management system.

Keywords: Internet of things, environmental sensor, device management, sensor configuration

1. Introduction

Recently, numerous Internet of things (IoT) environmental sensor devices are being installed in houses, buildings, and factories to measure surrounding environmental conditions such as temperature, humidity, illumination, CO_2, and air quality. For efficient energy management, the data measured from these environmental sensors are used to control related equipment such as lighting and air conditioners in real time.

However, different configurations should be applied to each environmental sensor depending on the characteristics of the place where the sensor is installed and the purpose of installation. Also, the data storage to which environmental sensors send data is different for privacy and corporate information security. For these reasons, there are limitations in collectively applying the same device image to each environmental sensor. In addition, there are difficulties in maintenance of installed devices. In this paper, to separate the device software image and sensor configuration information, we introduce an IoT environmental sensor management scheme based on sensor configurations in conjunction with a software configuration management (SCM) system.

2. Sensor Configuration-based Management

2.1 Sensor Configuration

In our previous work, we developed an IoT environmental sensor module that connects to the Raspberry Pi, an open hardware platform. This sensor module measures temperature, humidity, CO_2, illuminance, sound, total volatile organic compounds, and motion [1]. Each IoT environmental sensor device has device interface, sensor setting, and management configurations.

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (No. 20202020900290).

Fig. 1. Sensor configurations

This configuration information varies depending on the purpose, location, environment, and management organization of the sensor device. The recent trend for sensor management is to collect and manage both sensor identification and management information in a centralized Cloud service. However, to meet the requirements of the privacy protection and corporate information security, this paper proposes a device management scheme based on sensor configurations with a dedicated Git server, a SCM system.

Sensor configurations are defined in JSON format as shown in **Fig. 1**. They are composed of three parts: node, sensor, and server. The node part has a unique device identifier, device configuration information such as a HDMI port, LEDs, display status, and software management configurations. The sensor part includes the activation states of each environmental sensor and the BLE scanner for indoor localization based on Bluetooth beacon signal strength. The server part specifies the server address and protocol to which sensing data will be transmitted. Currently, MQTT, Modbus, Apache Kafka, IBFRAME Ingress Router [2], HTTP/TCP/UDP server protocols are supported.

2.2 Device Configuration Management

Since the sensor configuration is completely separated from the sensor device software image, all devices with the same environmental sensor module have the same software image installed. To deploy a sensor on a new site as shown in **Fig. 2**, we create a new JSON sensor configuration

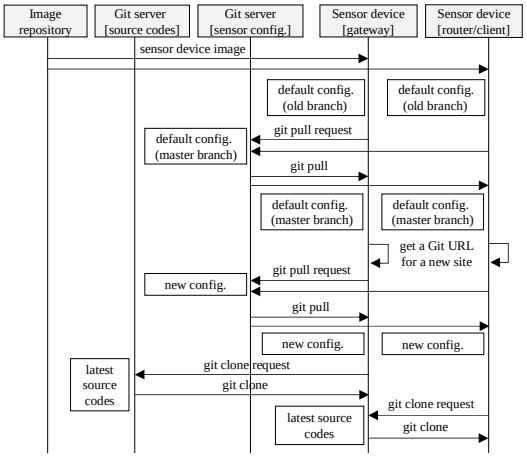


Fig. 2. Sensor device management

for the site and upload it to Git server. Then we update the Git URL of the newly configured JSON file to the master branch of the default sensor configuration in the Git server. The common image of the IoT environmental sensor device basically has the old branch of the default sensor configuration and is automatically updated with the master branch of default sensor configuration after first booting. The sensor device applies the new sensor configuration pointed to by the master branch, and the sensor configuration process is finished after updating the sensor device software.

3. Conclusions

This paper proposed the SCM system-based device configuration management scheme and sensor configuration structure to efficiently manage and maintain IoT environmental sensor devices installed in multiple sites.

References

[1] Y. Ji, K. Ok, and D. Kwon, "The Indoor Environment Monitoring System for Intelligent Buildings Using Wifi Mesh Based Internet of Things," in *Proc. of ACM International Conference on Mobile Systems, Applications, and Services (MobiSys)*, 2020.

[2] D. Kwon, K. Ok, and Y. Ji, "IBFRAME: IoT Data Processing Framework for Intelligent Building Management," in *Proc. of International Workshop on IoT Big Data and Blockchain (IoTBB) in conjunction with IEEE International Conference on Big Data (IEEE Big Data)*, 2019.

VR-Based Simulation System for the Coriolis Illusion During Flight

YoungBeom Kim and Kwang-Mo Jung
Korea Electronics Technology Institute, Korea
[e-mail: ybkim@keti.re.kr, jungkm@keti.re.kr]
*Corresponding author: Kwang-Mo Jung

Abstract

This paper presents the VR-based simulation system for training of the Coriolis illusion during flight. The simulation system is composed of a display system, a simulation system, an input system, a motion system, and a training scenario. The display system shows the simulated scene to a pilot and the simulation system generates flight data by using flight dynamics and user controls. The motion system generates the movement of the motion platform and the input system receive pilot control data and transfer to the simulation system. The training scenario contains flight parameters to generate the Coriolis illusion effect. The proposed system offers the authentic flight experience which includes the Coriolis illusion effect and makes it possible to a low-cost, but high-quality training of spatial disorientation.

***Keywords*:** Virtual reality, flight simulation, virtual training, motion platform

1. Introduction

Virtual Reality (VR) has enlarged the area from traditional applications such as games and 360 videos to the immersive training systems. The Flight simulator industry also regards the VR-based simulator as the future of pilot training [1].

The proposed system focus on the Coriolis illusion which is one of the top seven most commonly experienced illusions during flight [2]. The Coriolis illusion effect has occurred when prolonged constant-rate turn. During the constant-rate turn, the pilot's motion-sensing system can not sense the turning, because of the limits of the human's vestibular system. In this disoriented condition, abrupt head movements can create the illusion of movement in an unexpected axis and lead the fatal mishap.

The proposed system enables us to experience a dangerous situation at a low-cost by using VR technologies. Traditional flight simulators [3] require a huge and expensive motion platform and projection display. However, our system only requires a commodity low-cost motion platform and VR HMD devices.

2. System Structure

2.1 Over-view

The proposed simulator system is composed of the four systems and a training scenario, as shown in **Fig. 1**. In this chapter, each of the systems and the scenario are described.

2.2 The Display System

The display system shows the simulated scene to pilots. Pilots see the scene through a light-weight

This work is supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 20DRMS-B147026-03).

Fig. 1. System flow

wireless head-mounted display. As a result, the proposed system does not require a huge projection screen and decreases the pay-load of motion platforms.

2.3 The Simulation System

The simulation system simulates flight data by using user control data. Flight data contains the pose of the airplane and the scene from the viewpoint of the pilot. The pose of the airplane is computed by using aerodynamics and is sent to the motion system for generating movement of the motion platform. The rendered scene is shown to a pilot through the display system.

2.4 The Motion System

The motion system generates the actual movement of the motion platform. Simulators for the Coriolis illusion experience require a motion platform that supports infinity yaw motion, because, the Coriolis illusion occurs in the prolonged constant-rate turn. As a result, the traditional simulator requires a huge motion platform to endure high payloads that arise from heavy display systems. However, the proposed system has a light-weight display system and it makes possible the commodity motion platform to generate the Coriolis illusion experience.

2.5 The Input System

The input system receives a pilot's control data. The controllers of this system are composed of three components, that are a yoke, a throttle quadrant, and a flight rudder pedals. The components are attached to the shell of the motion platform for free movements.

2.6 The Training Scenario

The training scenario describes flight parameters for implementing the Coriolis illusion effect. The important parameters are an angular velocity on the yaw-axis and flight routes for prolonged turns, but other parameters such as the ToD (Time of Day) and weather conditions are also included.

Fig. 2. Simulator operation result

3. Conclusions

This paper presents the VR-based simulation system which enables the authentic flight experience with the Coriolis illusion effect. The system makes it possible to a low-cost, but high-quality training about the Coriolis illusion during flight. The experiment and pilot test for obtaining the parameters such as angular velocity on the yaw-axis will be the future work.

References

[1] R. J. Poisson and M. Miller, "Spatial Disorientation Mishap Trends in the U.S. Air Force 1993 – 2013," *Aviation Space and Environmental Medicine*, , Vol 85, No. 9, pp. 919-924, 2014.

[2] R.S.J. Matthews, F. Previc, and A. Bunting, "USAF Spatial Disorientation Survey," in *Proc. of RTO Human Factors and Medicine Panel (HFM) Symposium*, 2002.

[3] D. Klyde, A. Lampton and P.C. Schulze, "Development of Spatial Disorientation Demonstration Scenarios for Commercial Pilot Training," in *Proc. of AIAA Modeling and Simulation Technologies Conference*, 2016.

Training Monitoring System for Spatial Disorientation Simulators

YoungBeom Kim and Kwang-Mo Jung
Korea Electronics Technology Institute, Korea
[e-mail: ybkim@keti.re.kr, jungkm@keti.re.kr]
*Corresponding author: Kwang-Mo Jung

Abstract

This paper presents a training monitoring system for spatial disorientation simulators. The system shows flight information, bio-data of a trainee, and a spatial disoriented score, in real-time. The system is composed of four subsystems. The flight data monitoring subsystem is plugged into traditional simulation programs and sends simulated data to the data management subsystem. The bio-data monitoring subsystem gathers bio-data such as gaze, EEG (electroencephalography), and ECG (electrocardiogram) and sends the data to the data management subsystem. The data management subsystem records the training data in real-time. The display subsystem loads the gathered data and computes the spatial disoriented score and then shows them to the trainer in instructor operator stations. The proposed system enables us to the efficient analysis of the spatial disorientation training as supporting the visualization of related data in a display.

***Keywords*:** Virtual training, spatial disorientation, bio-data, flight simulator, monitoring system

1. Introduction

The spatial disorientation (SD) is the status of a pilot that misinterprets the pose of aircraft. The misinterpreted pose, which includes attitude, altitude, or airspeed, leads to mishaps and the SD related mishap is the second greatest factor of fatalities [1].

The proposed system focuses on obtaining the training data and computing the SD score of pilots. Recently, the relationship between visually induced SD and EEG are analyzed [2]. In this paper, the SD score is computed by using gaze, flight data, and EEG data. The score can give a direct message to the trainer who does not an expert in the biomedical area.

2. System Configuration

2.1 Over-view

The proposed system is composed of four subsystems. In this chapter, each of the subsystems is described.

2.2 The flight data monitoring subsystem

The flight data monitoring subsystem tracks the flight data from the simulator. The data includes the instruments' data such as the airspeed indicator (ASI), altimeter, vertical speed indicator (VSI), attitude indicator (AI), heading indicator (HI), and turn coordinator (TC). The flight route is also included. The tracked data is sent to the data management subsystem via calling Web API.

This work is supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 20DRMS-B147026-03).

2.2 The bio-data monitoring subsystem

The bio-data monitoring subsystem tracks the gaze, EEG, and ECG as depicted in **Fig. 1**. The heart-rate and stress score is computed by using the raw signal of ECG. The attention scores are analyzed from the EEG, which is tracked by using a 14-channels tracking device. The whole data is transferred to the data management subsystem.

Fig. 1. The screenshot of the bio-data monitoring subsystem

2.3 The data management subsystem

The data management subsystem has a database and operated as a web server. Both of the data monitoring subsystems call Web API, and then the data management subsystem interprets the API call and records the data to the database.

2.4 The display subsystem

The display subsystem shows the tracked data and the SD score as shown in **Fig. 2**. The ECG, EEG, and HR data are shown on the right side. The center panel presents gaze data. Note that the configuration of data can be customized.

Fig. 2. The screenshot of the proposed system.

3. Computation of the SD score

The SD score presents the probability that the pilot is disoriented. The pose of the simulated airplane and the disoriented pose information, which is described in an SD training scenario, are compared and if the difference is under threshold then the score is linearly increased. If this situation is sustained for 30 seconds, the value has the maximum value. If an elapsed time between the present and the last attitude check time is longer than the threshold or the increased alpha value is observed for lead P4, which in the right frontal lobe, then the increasing slope is doubled by adapting the recent experiment result [2].

4. Conclusions

In this paper, we present a training monitoring system for spatial disorientation simulators. Future research will be in direction of parameter tuning for computation of the SD score.

References

[1] N. Michael, L. Ben, R. Angus and M. Braden, "The Role of Perceptual Modeling in the Understanding of Spatial Disorientation During Flight and Ground-based Simulator Training," in *Proc. of AIAA Modeling and Simulation Technologies Conference*, 2012.

[2] M. Bauer, B. Bruegge, G. Klinker et al., "EEG Analysis of Visually Induced Spatial Disorientation," in *Proc. of 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2019.

An automatic generation of scene description for video retrieval by Korean image captioning and face recognition

Choongsang Cho, Young Han Lee, Min Soo Ko, and Hyok Song
Intelligent Image Processing Search Center, Korea Electronics Technology Institute
Gyeonggi-do, South Korea
[e-mail: {ideafisher, yhlee, kmsqwet, hsong}@keti.re.kr]
*Corresponding author: Choongsang Cho

Abstract

This paper presents an automatic generation of scene description tags for video retrieval by using Korean image captioning and face recognition methods. The proposed system provides a visual scene description that explains the global scene with a sentence and a human identification through the face recognition scheme. For this work, a Korean image captioning method is developed and trained by using the Korean MS-COCO image captioning dataset and targeted dataset for a specific service domain. Also, a face recognition scheme is developed with a face dataset of celebrities in the specific service domain, and face region for the recognition is obtained by using a pretrained face detection model with a public train dataset. Those results are used as scene description tags for video retrieval, and the methods to generate visual scene description tags are implemented into a single system. Therefore, we believe that the proposed scheme can be a useful approach for video and image retrieval based on scene understanding.

***Keywords*:** Scene description, Korean image captioning, face recognition, video retrieval

1. Introduction

Over the past few years, computer vision technology has significant progression based on deep convolutional networks [1-4]. In various studies with deep learning, a creation of the natural descriptions of images has been actively conducted [2-3]. In addition, automatic face localization that is the prerequisite step of facial image analysis has been widely studied [4]. The individual information from those methods is not enough to explain the situation in the image.

Therefore, we proposed a system that provides a visual scene description with a sentence and a human identification, and that is applied those schemes into video analysis and retrieval system as shown in **Fig. 1**.

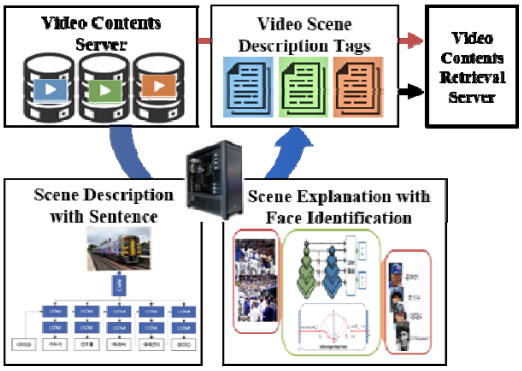


Fig. 1. An automatic scene description generation with Korean sentence and human identification.

This research was supported by a research grant from MSIT/IITP [2020-0-01982]

Fig. 2. Visual scene description results by proposed methods and system

2. Proposed Method and Implementation

The architecture is illustrated in **Fig. 1**. It consists of a generation of scene descriptions in a sentence and human identification by face recognition. To generate the scene explanation by sentence, the Korean image captioning method is developed and trained by using the Korean MS-COCO image captioning dataset and targeted dataset that has been established for specific sports. In addition, a face recognition scheme is developed by comparing the face features of celebrities in the specific service domain with an addictive angular margin penalty. To extract the features, a localized face region is obtained by using a pretrained face detection model with a public train dataset. The extracted description of main video scenes is transmitted to the video analysis and retrieval system, and it is used as a tag of those scenes, as shown in **Fig. 1**.

3. Experiment

In order to implement the automatic generation of multiple video descriptions, two deep learning methods, Korean image captioning and face recognition, are deployed into a single system.
The selected images that are the main scene in the video are applied to the proposed method, and then the results are shown in **Fig. 2**. As illustrated in the results, the scene is described by the Korean sentence and human name when the face is a registered person, otherwise represented by the sentence and a symbol 'UnKnown'. The proposed methods are integrated with an analysis and retrieval system, as shown in **Fig. 3**. The analyzed video is represented by the description and searched by the words or sentences.

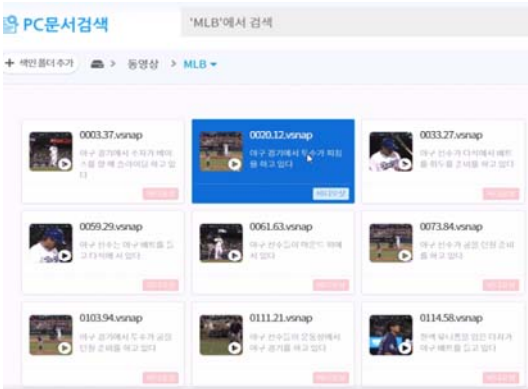


Fig. 3. Video analysis and retrieval system integrated with proposed methods

4. Conclusions

This paper presents an automatic scene description generator for video retrieval by developing Korean image captioning and face recognition. Those methods provide visual scene explanation with a sentence and human identification, and the results are very useful for video analysis and retrieval.

References

[1] Vinyals, O., Toshev, A., Bengio, S., Erhan, D.: "Show and tell: A neural image caption generator," In: proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3156-3164, 2015.

[2] DAI, Bo; LIN, Dahua. "Contrastive learning for image captioning," In Proc. of Neural Information Processing Systems, pp. 898-907, 2017.

[3] LU, Jiasen, et al. "Knowing when to look: Adaptive attention via a visual sentinel for image captioning," In: Proc. of the IEEE conference on computer vision and pattern recognition. pp. 375-383, 2017.

[4] DENG, Jiankang, et al. "RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild," In Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 5203-5212. 2020.

A real-time de-identification by object detection and elimination of human identification information from video contents

Min Soo Ko, Young Han Lee, Choongsang Cho, and Hyok Song
Intelligent Image Processing Search Center, Korea Electronics Technology Institute
Gyeonggi-do, South Korea
[e-mail: { kmsqwet, yhlee, ideafisher, hsong}@keti.re.kr]
*Corresponding author: Hyok Song

Abstract

This paper presents a real-time de-identification scheme for video content by using object detection and elimination methods. The proposed method consists of an object detection method with deep-learning and two elimination schemes to remove the identification region in the contents. For this work, the object detection model with the dataset including human identification; person, faces, identity (ID) cards, and license plates, is trained. Then, the detected identification of humans is de-identified through the mosaic or blur method. For real-time implementation, the object detection model and the elimination methods are optimized to reduce the computational time by using TensorRT and GPU (Graphic Processing Unit) computation, respectively. By the proposed scheme, it required 25 ms to eliminate the identification region in an image of contents. Therefore, we believe that the proposed scheme can be a useful method and system for a real-time de-identification system in broadcasting contents, etc.

***Keywords*:** De-Identification, object detection, elimination method, deep-learning

1. Introduction

Over the past few years, computer vision technology has significant progression based on deep convolutional networks and GPU (Graphic Processing Unit) [1-3]. In various studies with deep learning, object detection methods have been widely researched and applied for intelligent security and autonomous vehicle system [2]. In recently, various video contents is exponentially increased in many areas, therefore personal identification information in those contents should be automatically removed to cope with huge contents. De-identification that eliminate the identity information has been commonly used in various fields; communications, multimedia, social networks, and video surveillance [3]. We proposed a real-time de-identification system that eliminates the person, faces, identity (ID) cards, and license plates region in the video contents as shown in **Fig. 1**.

2. Proposed Method and Implementation

The proposed real-time de-identification architecture consists of an object detections

This research was supported by a research grant from MSIT/IITP [2020-0-01982]

method with deep-learning and two elimination schemes to remove the identification region in the contents. To obtain identity regions of humans in the contents, an object detection scheme is trained with the dataset including person, faces, ID cards, and license plates. To eliminate those regions, mosaic or blur methods are deployed to make it invisible. To reduce the computational time, the complexity of object detection is decreased by TensorRT that provides an automatic merge of deep-leaning layers, and data precision control. Also, the complexity of elimination methods is reduced the parallel processing by GPU.

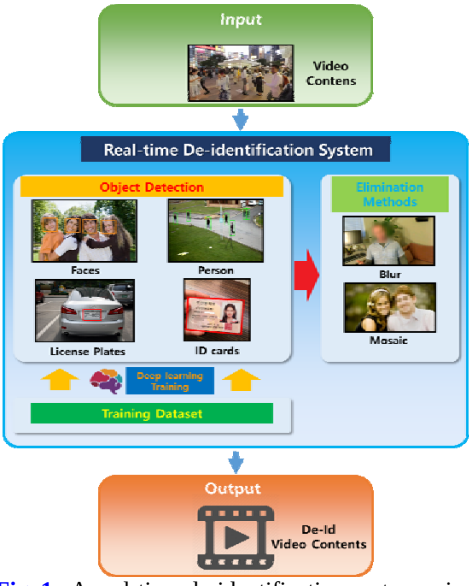


Fig. 1. A real-time de-identification system using object detection and elimination methods.

Fig. 2. Identity region detection results by the proposed method.

3. Experiment

For the experiment, video surveillance contents are used, and four identity regions; person, face, ID card, and license plate, are detected by the proposed system as shown in **Fig. 2**. Those regions are eliminated by the system as shown in **Fig. 3**.

4. Conclusions

In this paper, we present a real-time de-identification scheme for video content by using object detection and elimination methods. To implement the real-time system, the complexity of object detection and elimination is reduced by applying TensorRT and parallelizing elimination methods on GPU. By the proposed scheme, it required 25 ms to eliminate the identification region in an image of contents.

Fig. 3. Identity elimination results by the proposed methods; (left: blur, right: mosaic).

References

[1] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao: “Yolov4: Optimal speed and accuracy of object detection,” arXiv preprint arXiv:2004.10934, 2020.

[2] Sreenu, G., Durai, M. : “Intelligent video surveillance: A review through deep learning techniques for crowd analysis,” In: proc. of J. Big Data, pp. 48–75, 2019.

[3] Z. Ren, Y. J. Lee, and M. S. Ryoo: “Learning to anonymize faces for privacy preserving action detection,” In: proc. of European Conference on Computer Vision, 2018.

Entry Prediction into Hazardous Areas with Encoding Crowd Interaction

Inkyu Choi, Min Soo Ko, Young Han Lee, and Hyok Song
Intelligent Image Processing Search Center, Korea Electronics Technology Institute
Gyeonggi-do, South Korea
[e-mail: { cig2982, kmsqwet, yhlee, hsong}@keti.re.kr]
*Corresponding author: Hyok Song

Abstract

Many accidents and casualties have occurred in hazardous areas such as construction and industrial sites. To cope with these problems, technology to predict the risks in advance is critical for human safety. We propose a method to prevent accidents in advance by determining whether a target enters a dangerous area using a pedestrian trajectory prediction scheme. Encoded crowd interactions and estimated future trajectories by a regression method are obtained through a deep-learning model. For our study, the exit regions of the Grand Central (GC) that is a public dataset is defined as hazardous areas for training and evaluation of the trajectory prediction model. In the evaluation, the entry prediction into the hazardous area has an accuracy of 92.7% by the proposed method. Therefore, we believe that the proposed method is useful tools to prevent safety and security accidents.

***Keywords*:** trajectory prediction, crowd interaction, hazardous areas , deep-learning

1. Introduction

As deep learning technology has developed and its excellence has been proven, it has begun to be applied to various application fields such as medical, economics, education, automobiles, and culture. Industrial accidents must be prevented because they cause not only material damage but also human damage. In industrial sites, human damage is caused by the carelessness of workers, and the best solution is to predict what will happen in the future. Trajectory prediction is a technique that predicts the future location by analyzing the past path of an object. In this paper, to prevent industrial accidents, we propose a method of predicting whether a person enters a dangerous zone using trajectory prediction. Recently, GAN-based trajectory prediction[1,2] has been studied, but in this paper, we use an LSTM-based technique[3] that predicts one trajectory per person.

2. Proposed Method

2.1 Trajectory prediction model

The Trajectory prediction model consists of three modules: motion encoder, MLP location encoder, and displacement prediction. The motion encoder module is designed to represent different individual movement patterns. Since LSTM is very useful for processing time-series data, it uses it to encode the movement of each pedestrian.

The location encoder module is designed to represent the positional relationship between the prediction target and its neighbors. It consists of a multi-layer perceptron containing three layers with ReLU activation function. After extracting the encoded vector using the coordinates of each person's current location as input, spatial affinity is calculated through the inner-product between

This material is based upon work supported by the Ministry of Trade, Industry & Energy(MOTIE, Korea) under Industrial Technology Innovation Program(20007872)

the vectors. Then, a crowd interaction vector is generated by the sum of product operation between the motion vector extracted from the motion encoder and the spatial affinity calculated from the location encoder.

The displacement prediction module is responsible for estimating the location displacement of each person between the present and the future. It is composed of one fully connected layer and acquires location displacement by regression.

2.2 Defining Hazardous Areas

In the Grand Central (GC) scene, as shown in **Fig. 1**, the entrance or escalator boarding gate with a large floating population was defined as a Hazardous Areas. (Red ellipse area)

Fig. 1. Hazardous areas definition in Grand Central

3. Experiments

In our experiments, we have observed the trajectory for 5 frames and use them to predict the trajectory for the next 5 frames. It evaluates whether the trajectory predicted by the deep learning model is located within the hazardous area by inputting only observed trajectory data that the ground truth trajectory belongs to in the dangerous area. Figure 2 is a visualization of the predicted trajectory for an arbitrary person. Green is the observed trajectory, red is the ground truth, and blue is the predicted trajectory. As a result of the test, the proposed method predicted whether to enter the danger zone with 92.7% accuracy.

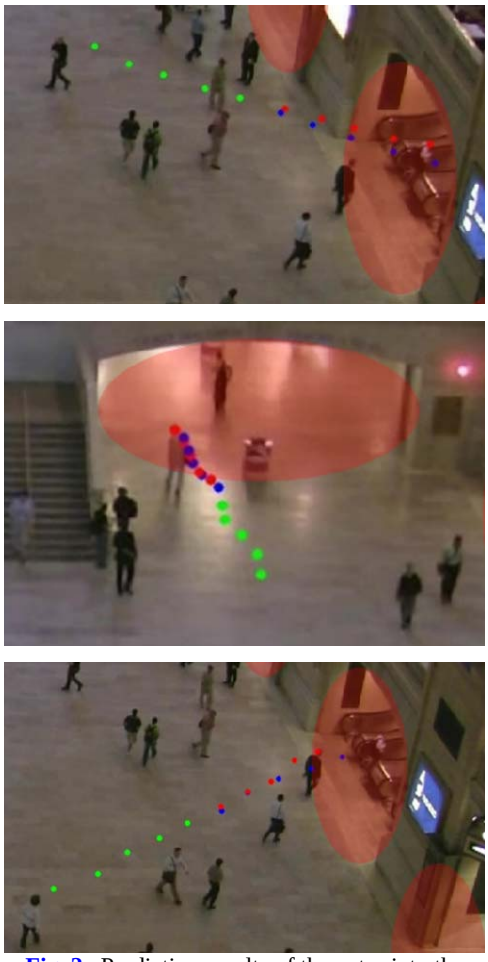


Fig. 2. Prediction results of the entry into the hazardous areas

4. Conclusions

In this paper, we propose a method of predicting in advance whether people enter hazardous areas using trajectory prediction. The trajectory prediction technique considers the different importance of all pedestrians based on their spatial affinity to the target pedestrian. In the grand central, we defined areas with a large floating population as dangerous areas and evaluated whether the trajectory predicted through the deep learning model was included in the dangerous areas. As a result of the experiment, we confirmed that the proposed method classifies whether people enter hazardous areas with an accuracy of 92.7% in real-time.

References

[1] A. Gupta, J. Johnson, L. Fei-Fei, S. Savarese and A. Alahi, "Social gan: Socially acceptable trajectories with generative adversarial networks" *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2255-2264, 2018.

[2] J. Amirian, J. B. Hayet and J. Pettré, "Social ways: Learning multi-modal distributions of pedestrian trajectories with gans" *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019.

[3] Y. Xu, Z. Piao and S. Gao, "Encoding crowd interaction with deep neural network for pedestrian trajectory prediction" *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5275-5284, 2018.

KSII The 12th International Conference on Internet (ICONI) 2020.

Design and Implementation BMC based Edge Server System Control Module in Rugged Environment

Seung Gwon Kim, Younghwan Kim, Jae-Hoon An
Human IT Convergence Center, Korea Electronic Technology Institute
Seongnam-Si Gyeounggi-do, South Korea
[e-mail: 09labs@keti.re.kr, yhkim93@keti.re.kr, corehun@keti.re.kr]
*Corresponding author: Jae-Hoon An

Abstract

Edge computing technology is used to reduce data analysis processing delays in areas such as smart cities, smart factories and autonomous driving. However, present edge computing technology uses hardware that is unsuitable for rugged environments, i.e. industrial area. As a result, it cannot be applicate to the closest environment to smart devices, there is a limit to reducing data analysis and processing delays. Therefore, it is necessary to develop edge computing technology that can be applied in rugged environment. In this paper, we proposed BMC (Baseboard Management Controller) based edge server system control module that can be applied in rugged environment.

***Keywords*:** Edge Server, Edge Computing, BMC, IPMI, Redfish, Rugged Environment,

1. Introduction

The Centralized server computing, storage and networking, central data centers are face problems such as data processing delays and performance bottleneck [1]. Due to this problems, the demand for edge computing technology is increasing. Edge computing technology enables faster computation and data transmission because the phisical distance between the user and the edge system is short and the amount of data is less than cloud computing [2]. In order to reduce the physical distance between the device that generates data and the server, an edge server system applicable to the site is needed. However, the tolerance temperature of edge server solution that provided by existing server manufacturer is 0 to 55 celsus. So it cannot be applied to rugged environment because there is a limit to reducing the physical distance between edge device and edge servers. Therefor we need to consider to develop edge computing technology that can be applied in rugged environment. In this paper, we proposed BMC based edge server system control module that can be applied in rugged environment.

2. BMC based Edge Server Control

It is very important to monitor the chassis sensors and control the chassis temperature to implement an edge server system that can be applied in rugged environment. Therefore, the SoC (System on Chip) is required for sensor monitoring and control independently of the edge server system. BMC is installed on the motherboard to communicates with the system manager through an independent connection. And it plays the role of controlling and monitoring the overall server.

2.1 System Design

The interface is required for local or remote administrators to access the BMC. Basically, administrators use the IPMI (Intelligent Platform Management Interface). The main feature of IPMI include sensor and system monitoring,

recovery, abnormal event logging and server remote control. It is universal server remote control interface, but it has limitation of low security, incompatibility of OEM commands, and scalability. In this research, we designed and the three elements of Edge BMC: (1) IPMI/DCMI Provider Library, (2) Hardware device driver and selection for use in rugged environments, (3) GUI based edge server control and monitoring interface.

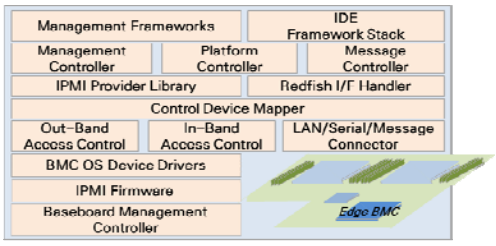


Fig. 1. Proposed Edge BMC Block Diagram

As a shown Figure 1, the block diagram of the proposed Edge BMC contains three elements defined in this study. The proposed Edge BMC includes a Device Mapper layer for controlling various peripherals such as chassis inlet and outlet sensors, heating module and fans based on AST2600 EVB, and includes a library capable at the same time supporting IPMI 2.0 and Redfish APIs. In this paper, we describe the system verification details to verify Edge BMC.

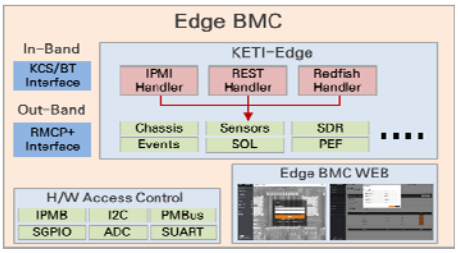


Fig. 2. KETI-Edge Process Block Diagram

First, as shown **Fig. 2**, a KETI-Edge process supports IPMI and Redfish API at the same time. It shares the resources that used for server monitoring and device control, and allows a large scale server management interface. Second, H/W access control block consists of device drivers such as PMBus, ADC and SGPIOs for various hardware controls. This allows the edge server system to run reliably in a rugged environment. Finally, Edge BMC Web supports a feature that users to easily monitor and remotely control the server in real time in a GUI interface other than ipmitool or redfishtool.

2.2 System Verification

In this research, we verified proposed Edge BMC system design using AST2600. It is suitable for rugged environment because tolerance temperature is -40 to 85 celsius. Currently, we implemented KETI-Edge process to monitoring and control server systems, and verified using ipmitool and Redfish API for compatibility. The Edge BMC firmware was verified by using a simulator similar to the rugged environment. Finally, we verified control and monitoring of edge server systems through Edge BMC Web.

3. Conclusions

In this paper, we show the feature and structure of Edge BMC for edge server management in a rugged environment. Through the design and implementation of the three feature of the Edge BMC, we showed the compatibility and scalability of the Edge BMC operable in the rugged environment. Based on the results of the design, we will continue to focus on large scale edge server system technology that can be applied to rugged environment through the Edge BMC.

4. Acknowledgments

This work was supported by **Institute for Information & communications Technology Promotion(IITP)** grant funded by the Korea government(MSIT) **(No.2020-0-00832, Development of Server Computing Module for Compute-intensive Edge Services)**

References

[1] CSM Babou, D Fall, “Hierarchical Load Balancing and Clustering Technique for Home Edge Computing”, Vol.8, IEEE Access, 2020.

[2] C.H.Hong and B.Varghese,“Resource management in fog/edge computing:A survey on architectures, infrastructure, and algorithms” ACM CSUR, vol. 52, no. 5, pp. 1-37, Sep, 2019.

Design and Implementation of Chassis Manager to control and Manage Edge Server in Rugged Environment

Young-Rim Choi , Jae-Hoon An, Young-Hwan Kim
Korea Electronics Technology Institute
Gyeonggi-do, South Korea
[e-mail: , dudfla1995@keti.re.kr , corehun@keti.re.kr, yhkim93@keti.re.kr]
*Corresponding author: Jae-Hoon An

Abstract

Due to Cloud computing technology has a limitation in processing data generated by IoT devices that have exploded, there is an increasing demand for edge computing technology that can solve data processing delay and transmission delay problems. However, the edge servers currently provided by manufacturers are difficult to use near industrial environment, considering various environmental factors at the industrial environment. When using an edge server in an industrial environment, the role of the chassis manager managing the edge server is important because of these environmental factors. However, since the existing chassis manager does not guarantee redundancy, there is a reliability problem that it is difficult to control and manage the edge server when the chassis manager is stopped. In this paper, we proposed a chassis manager that can monitor and control environmental factors to use an edge server in an industrial environment and guarantees operational reliability through redundancy.

***Keywords*:** Edge Server System , Rugged Environment, Monitoring, Redundancy

1. Introduction

With the spread of various IoT devices, the demand for edge computing is rapidly increasing[1]. Because of this, the demand for edge servers in industrial environments is increasing. For the normal operation of the edge server in an industrial environment, the role of the chassis manager who manages the edge server and the chassis enclosure is important. However, the chassis manager for edge servers produced by third-party vendors has an endurance temperature of -5 to 55°C. Considering the industrial environment such as outdoors, the existing chassis manager is not suitable. In addition, since most of the existing chassis managers are not in a redundant configuration, it is difficult to guarantee operational reliability of the edge server when the chassis manager is abnormally stopped. In this paper, we proposed a chassis manager capable of controlling and managing an edge server in a rugged environment. The proposed chassis manager secured operational reliability through hardware and software redundancy.

2. Chassis Manager Design and Implementation

In an industrial environment, the chassis manager must monitor and manage environmental factors and power conditions in real time to prevent abnormal stops. In this paper, we designed a system that satisfies the requirements of the Redfish API defined by DMTF. Redfish supports hardware and system software monitoring and control functions in a large-scale edge server system environment. Finally, in this study, the chassis manager designed to support network and data redundancy to ensure operational reliability.

2.1 Edge Server & Hardware Monitoring

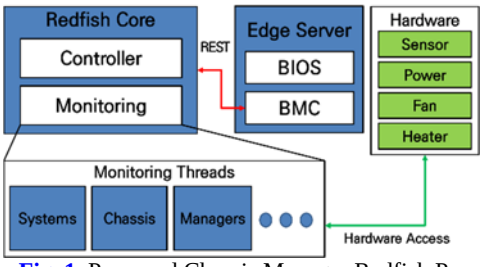


Fig. 1. Proposed Chassis Manager Redfish Process

The structural design of the chassis manager to satisfy the Redfish API requirements is shown in **Fig. 1**. The chassis manager controls the edge server through the Redfish Core Controller and provides a monitoring function. Monitoring targets were divided by Resource defined in Redfish API. Resources include Systems that manage Edge Servers, Chassis that manages chassis enclosures, and Managers that manages chassis manager remote access networks. According to Redfish API, Resource object can be extended and reconfigured.

2.2 Network & Data Redundancy

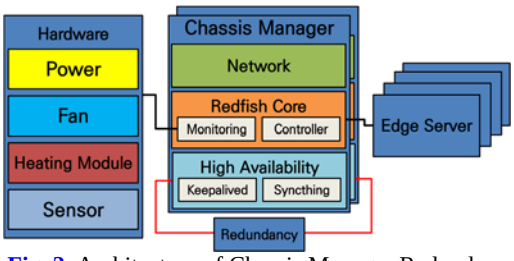


Fig. 2. Architecture of Chassis Manager Redundancy

In this paper, reliability was secured through hardware and software duplication of the chassis manager. As shown in **Fig. 2**, it is provided by the High Availability module. First, we implemented network redundancy with Keepalived. Two nodes (chassis manager) create public virtual IPs through Keepalived. Clients receive services through the virtual IP. There are two types of node operation: Active-Slave and Active-Active. Both methods ensure continuous service of the chassis manager. However, when the Active-Active method is used, both nodes provide services to reduce server load and enable low-latency services [2]. In this paper, data redundancy is implemented using Syncthing for real-time data synchronization. By using Syncthing, data consistency is guaranteed by synchronizing data between two nodes. By providing data consistency, operation reliability of the chassis manager is guaranteed.

2.3 System Verificatio

In this paper, we used Raspberry Pi 3 B+ as a chassis manager for system verification. We implemented the Redfish interface identically to two Raspberry Pis, and experimented with hardware monitoring and control through I2C connection with a hardware device. Also, for redundancy, we tested network redundancy with two Raspberry Pis through a virtual IP in the internal network. Data redundancy was tested data consistency between the two chassis managers using Syncthing.

3. Conclusions

In this paper, we propose a chassis manager for monitoring and controlling edge server systems that can operate in a rugged environment. Reliability is guaranteed through hardware and software redundancy of the chassis manager.

4. Acknowledgments

This work was supported by **Institute for Information & communications Technology Promotion(IITP)** grant funded by the Korea government(MSIT) **(No.2020-0-00832, Development of Server Computing Module for Compute-intensive Edge Services)**

References

[1] Djabir Abdeldjalil Chekired, Lyes Khoukhi, Hussein T. Mouftah, “Industrial IoT Data Scheduling based on Hierarchical Fog Computing: A key for Enabling Smart Factory”, IEEE Transaction on Industrial Informatics, pp 1, 2018

[2] Ashish Vulimiri, P.Brighten Godfrey, Radhika Mittal, et al, “Low Latency via Redundancy”, the ninth ACM conference, pp 283, 2014

Demonstration of disaster information and evacuation support services for the disabled and the elderly

Ilryong Kweon[1], Minho Son[1], Taeho Jung[2], Hanjun Lee[1], Onyu Kim[1], Yugyeong Kang[1]
PODO inc.[1], National Disaster Management Research Institute[2]
Seoul, South Korea
[e-mail: irk@ipodo.co.kr, mhson@ipodo.co.kr, thjung@korea.kr, hjlee@ipodo.co.kr, mhson@ipodo.co.kr, hjlee@ipodo.co.kr,]
*Corresponding author: Minho Son

Abstract

This study drew specific countermeasures for each crisis situation in consideration of the safety vulnerability characteristics in various disaster situations for the disabled and the elderly who are vulnerable to disasters. Based on this, the disabled and the elderly were investigated and analyzed for vulnerable facilities due to restrictions on movement, and the standards for alarm systems were designed to respond to more realistic disaster crisis situations, and pilot facilities were selected to minimize damage to the disabled and elderly living areas, and indoor space information of the facilities were established. In addition, optimal research results were derived through the design of disaster information and evacuation support systems through case analysis and the establishment of spatial information standards and data on indoor location measurement.

***Keywords*:** Disaster, evacuation map, hybrid, indoor positioning, VGI(Volunteered Geographic Information), 3D indoor map

1. Introduction

Every human being wants to live in a safe environment. However, they face various disasters such as unexpected natural and social disasters. Among them, social disasters in large-scale facilities such as rapid urban changes (such as new urban development and innovative cities) and the centralization of the population, the enlargement of facilities due to the development of construction technology, and the combination are increasing.

It also properly propagates disaster situations.

The number of disasters and damage caused by poor or poor evacuation is increasing every year.

Accordingly, the disaster management system is established to prepare for risks, but it is insufficient to respond to various social changes and demands.

Safety for the disabled, the elderly, children, etc. in overseas cases.

It is now in place to consider the vulnerable in the disaster preparedness stage.

In Korea, the need for a new paradigm for disaster management considering the safety-vulnerable classes (Framework Act on Disaster and Safety Management, 2018) such as the disabled, the elderly, and children emerges, and disaster management and safety management for the safety-vulnerable class.

Discussions are being made, but there is a lack of research and policy approaches considering their characteristics.

This study is part of a research project on the design of disaster information and evacuation support system for the elderly with disabilities supported by the National Institute of Disaster and Safety(Republic of Korea Government). Contract number 21193104700.

2. Related research

With the rapid change of climate and natural environment in modern society, Preparing for various disasters as the risk of disasters increases

The need and interest in managing the operation of disaster information to do so

It's increasing. Most of the disaster information built on it. The system is a non-disabled-oriented disaster information and management system.

For the safety of the vulnerable, such as the disabled, for disaster safety.

It is insufficient to deliver disaster information.

The disabled and the elderly are less cognizant in the event of a disaster. So they can safely evacuate or respond. Establishing a service system is urgent and disasters for the disabled and the elderly. Disaster Information Delivery and Evacuation Support System Considering Vulnerabilities. We need a system. To respond to disasters for the disabled and the elderly.

Customized disaster information service system was established. IoT-based integration for disaster information delivery and evacuation support system design for this purpose. Standardization of alarm systems by type of disaster based on control technology.

I'd like to present an evacuation assistance system.

3. Main subject

The purpose of this study is to suggest ways to effectively transmit disaster information and avoid disaster environment in case of disaster (fire) for the disabled and the elderly. The service aims to effectively share disaster situation information through visual services, hearing services, indoor location services, and user participation services. In particular, the service considering the user's situation considering the case where the disabled or the elderly can escape from the disaster environment by themselves (Out Bound) and when the movement is restricted or cannot be moved (In Bound- Rescue need). To provide. The concept of out bound and in bound is shown in **Fig. 2** below.

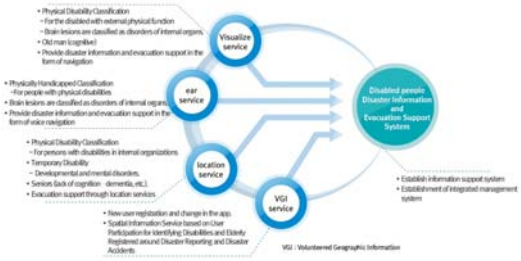


Fig. 1. Service architecture

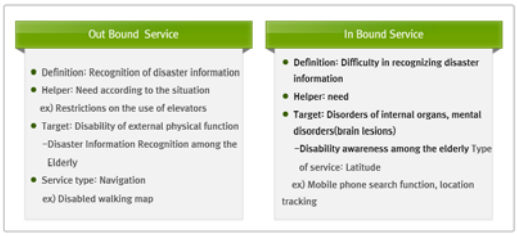


Fig. 2. Service Types according to Cognitive Ability of the Elderly with Disabilities

In order to share disaster situation quickly, this study intends to provide user participatory disaster sharing service. Most smartphone users have various SNS applications. Using these applications, we have configured a shared service in the cloud environment so that users can easily share disaster situations around.

Fig. 3. User Participation Disaster Information and Evacuation Support System 1) App-based disaster information and evacuation support services.
2) Spatial information based on user experience and participation(VGI).

By using this, the user will intuitively identify the location and utilize the geomagnetic field location technology to provide a path to avoid disasters in a short time.

4. Case study

This study proposes a disaster information service using a cloud-based app so that people with disabilities and the elderly can effectively identify disaster information and avoid disasters in disaster situations (fire). **Fig. 4.**

Fig. 4. Service flow chart

The Earth's magnetic field begins with the fact that the Earth is a big magnet. Any space on earth forms a certain level of magnetic field, and the strength of the geomagnetic field varies with location. In particular, the location of the interior space is distinguished by the characteristics of the geomagnetic field being distorted by steel structures, electronic equipment, etc. Unlike radio maps, there is also uncertainty, and there are many points similar to vector values measured at a particular location. The principle of location estimation was used by comparing the changes in the magnetic field and the pre-built magnetic field maps and patterns that occur when users move. In preparation for the worst-case disasters such as the power cut off, indoor measuring technology using geomagnetic fields was applied.

Although there are differences in measurement positions, overall, the error range decreases and accuracy increases as we move indoors. **Fig. 6.**

Cases by type of disability and type of activity are defined, and scenarios by type of disability (visual, auditory, and activity) and by type of activity (exploitation, rescue device, normal) are constructed as follows. **Fig. 7.**

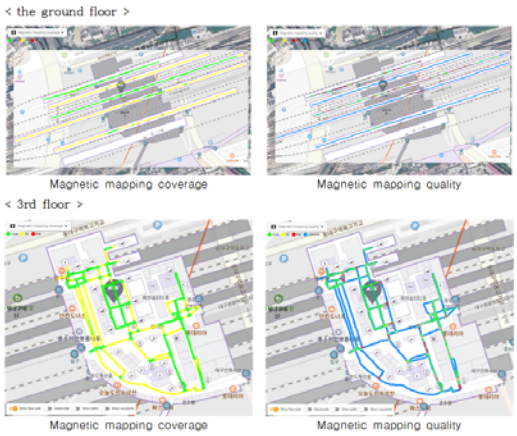


Fig. 5. Test Data with IndoorAtlas

5. Conclusions

The disabled are aware of the initial situation and respond to a disaster.

It is difficult and often leads to great damage. In fact, the number of deaths from fires per 100,000 people in Korea is 4.7 times higher than that of non-disabled people (Ministry of

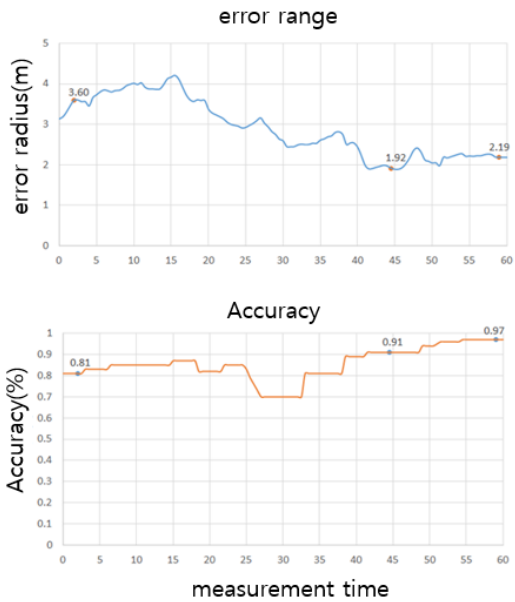


Fig. 6. Measurement of error range and accuracy of geomagnetic field maps

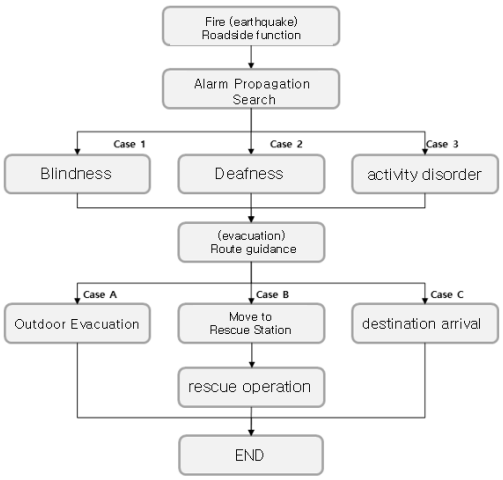


Fig. 7. Case by disability type and activity type

Public Administration and Security, 2017). Therefore, disaster safety management measures that reflect the characteristics of the disabled should be prepared and effectively responded to. Accordingly, the disaster information and evacuation support system for the elderly with disabilities were derived, but legal limitations such as sharing the location of users (Personal Information Protection Act) of indoor location measurement and theoretical limitations that have yet to be technically standardized were derived.

If continuous research is carried out due to technical limitations including technology, it can be a more realistic alternative to the safety-vulnerable class in disaster due to legal improvement and technological development, which is expected to make the use of indoor spatial information more active.

Fig. 8. Comparison with existing geolocation services

References

[1] National Institute for Disaster and Safety report "Disaster Information and Disaster Evacuation System Design for the Disabled and the Elderly" report, 2019.

[2] Korea Development Institute for Persons with Disabilities report "Development Manual of Disaster Response Behaviors Customized for Irum Center", 2018.

Real-Time Automobile Type and License Plate Recognition System

Jeong-Ah Kim, Ju-Yeong Sung, Se-Ho Park
Korea Electronics Technology Institute, Contents Convergence Research Center
Seoul, South Korea
[e-mail: kpkk1518@gmail.com, wuddl@keti.re.kr, sehopark@keti.re.kr]

Abstract

This paper studies a method to recognize vehicle types and license plates based on the Deep Learning model. We trained each algorithm through an automobile training dataset and analyzed the performance to determine what is the optimized model for vehicle type recognition. We trained and tested YOLO, SSD, and Faster-RCNN models. The YOLOv4 model outperforms other methods, showing 93% accuracy in recognizing the vehicle model.
Also, we introduce a real-time Automatic License Plate Recognition system using Deep Learning Models. We made the YOLOv4 model computationally lighter by eliminating the ROI setting step, without deteriorating recognition performance. Conventional license plate recognition systems exhibit two main problems. First, clear license plate visibility is required. Second, processing actual field data is computationally intensive and the ROI needs to be set. To overcome these problems, we performed plate localization directly on the entire image and conducted research taking low-quality license plate detection into account.
We aim to recognize the license plates of cars moving at high speeds on the road as well as stationary cars using the NVIDIA Jetson TX2 module, which is an embedded computing device.

***Keywords*:** object detection, deep learning, automobile type recognition, automobile license plate recognition.

1. Introduction

As the social structure becomes complex and the new transportation system emerges, various types of vehicles appeared on roads and parking lots. Each vehicle can be classified by the type of vehicle or capacity of the vehicle. For example, it can be classified as a passenger car and freight car or light car and a medium-sized car. Road driving fees and parking fees are different depending on the vehicle type, in order to reduce possible environmental pollution and to efficiently use road resources. If automatic vehicle type recognition is possible, collecting road fees or parking management fees might be convenient. In addition, 24-hour continuous monitoring is possible without human resources, and the overload of the manager can be reduced. Therefore, we decided to develop an object recognition system that distinguishes vehicle types.

There are many license plate recognition systems these days. Most of these systems target large or clear license plates. However, license plate data collected from the actual field are not always appropriate for processing. License plate images comprising field data are often too small or unclear. In fact, field data is often blurred because vehicles are moving at high speeds or there is a lot of noise. Therefore, most of the

previous systems proposed have difficulty in processing field data. Recently, studies overcoming this problem have been published, but due to a large amount of computation, they cannot be used in real-time on light, embedded devices.

In the current field of computer vision, the object recognition model based on deep learning shows high accuracy, and various studies are being performed in object detection based on deep learning.

In this study, various deep learning-based object recognition models were analyzed, and the performance was evaluated by the accuracy and processing speed. Our purpose is to create a system that allows real-time automobile type and license plate recognition on a relatively light device with less computation. To do this, we need to develop a way to recognize relatively small license plates from high-resolution images, such as 3K or 4K. This allows you to skip the step of finding the car area, reducing the amount of computation, and speeding up the system. Create the optimum model for recognizing the automobile type by comparing the deep learning model, and perform vehicle type recognition and license plate recognition in parallel.

2. Object Detection Technology based on Deep Learning

2.1 Object Recognition Technology

Object recognition technology that determines the presence of an object in an image is the ultimate goal of computer vision and machine learning. In previous object recognition research, the features of objects are selected based on human knowledge, and objects are detected based on the features.

Recently, as the performance of GPUs has improved and vast amounts of data can be collected through the Internet, object recognition technology based on a Convolutional Neural Network (CNN)[1] is drawing attention. The object recognition methods using deep learning surpass the existing algorithms based on human knowledge in performance and become mainstream in this field. Studies on object recognition based on deep learning have been actively conducted, and many models with better performance have emerged. Technologies for recognizing multiple objects in images have been developed. Also, studies are ongoing to develop advanced technologies recognizing objects by only part of them. There are several representative deep learning-based object recognition technologies including R-CNN [2], Fast R-CNN [3], Faster R-CNN [4], YOLO [5], and SSD [6].

2.2 Faster R-CNN

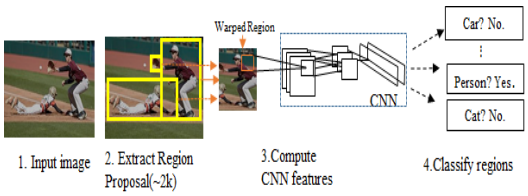


Fig. 1. Object Recognition Sequence of R-CNN

R-CNN detector is composed of two stages, regional proposal, and classification. First, through Selective Search, the detector finds 2,000 boxes that show the region of the target object. Then, classification is performed by applying CNN to all Bounding Boxes. As the amount of computation cost increases and the processing speed becomes slow. To compensate for the processing speed, Faster R-CNN performs object detection once in the output feature map after passing through CNN. Faster R-CNN uses the Region Proposal Network to solve the bottleneck caused by the selective search algorithm. Comparing the processing speed, Faster R-CNN is 200 times faster than R-CNN. The object recognition method of the R-CNN series has the disadvantage of slow processing speed, not suitable for real-time applications.

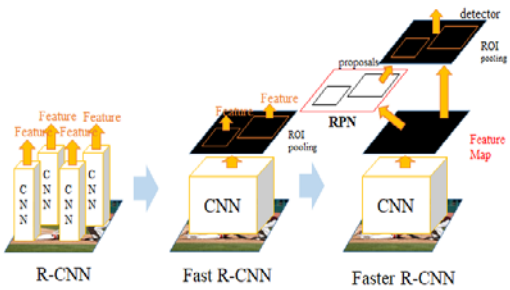


Fig. 2. R-CNN series structure

2.3 You Look Only Once

YOLO (You Look Only Once) and SSD (Single Shot Detector) are 1-stage-detectors, as Regional Proposal and Classification are simultaneously performed. YOLO was implemented in a way similar to the human object recognition system. The front part of the YOLO network structure is a modified structure of GoogleNet. In general, the deeper the Convolution Neural Network, the better the performance with more layers. However, as the network gets deeper, the number of parameters to be learned increases. The problem of the deep convolution operation is a huge amount of calculation and a huge number of parameter values that must be set. Therefore, Network is used to express the nonlinear relationship of data by using Multi-Layer Perceptron in convolution operation. YOLO divides the image into S*S grid regions and predicts the bounding box of each grid region. Compute confidence represents the reliability of the box. At the same time, it predicts the class probability. After that, the box and class probability are combined and the object is found by using non-maximum suppression.

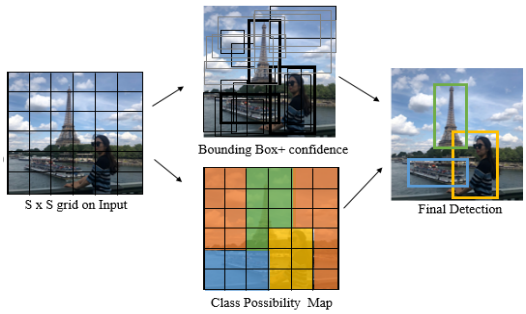


Fig 3. YOLO object detection

2.4 Single Shot Detector

Single Shot Detector[6] uses VGG16 as the base network due to its high-quality image classification, and then convolutional feature layers that progressively reduce in size are added to its end so that it is able to predict object at different scales by the aspect ratio. The core of SSD is predicting category scores and box offsets for a fixed set of default bounding boxes using small convolutional filters applied to feature maps. SSD uses the convolutional feature map at the front, It makes more detailed features. Moreover, performance and speed were improved by replacing the Fully Connected Layer.

3. Automobile-Type Recognition Approach.

3.1 Data Set for Automobile Type Recognition (KETI-ATR Dataset)

The vehicle type classification for the vehicle type recognition was based on the classification of the Korea Expressway Corporation. KETI-ATR dataset was divided into Light car (Labeled as compact), 9 passengers or less (Labeled as car), 9 passengers or more and 25 passengers or less (Labeled as mini_van), and 25 passengers or more (Labeled as big_van). Freight cars are classified into small freight cars (Labeled as mini_truck) if they are two-axle vehicles, large freight cars (Labeled as truck) for three or more-axle vehicles, and special freight cars. The data set for the experiment is shown in **Table 1**.

Table 1. Data for automobile type recognition

Labeling Name	TRAIN set	TEST set
car	1447	276
mini_van	244	55
big_van	33	8
mini_truck	516	163
truck	94	27
compact	286	39

YOLO, SSD, Faster-RCNN were trained with the same data set. We implemented YOLO using the recently released YOLO v4[7] to improve performance by larger resolution than before. The ROI of the data was extracted from the front window of the car to the bumper so that the features were not lost. An example of ROI extraction is shown in **Fig. 4**.

Fig 4. ROI designated area on automobile

3.2 Experimental Results of Automobile Type Recognition

To compare the result of vehicle type recognition by different deep learning-based models, the performance was evaluated by each method. After training the model, we chose the best model with the fastest processing speed and the highest accuracy. The test was performed with 450 images which are different from the training set, and the algorithms were implemented by GeForce RTX 2080ti. We measured the processing speed in Frame Per Second (FPS). Since the amount of data is not uniform, F1-score was measured together to evaluate the performance.

Fig. 5. Detected Automobile Image By SSD

In the case of the R-CNN model, since CNN is used and detection is performed in two steps, the accuracy is relatively high, while the speed is significantly slow. It might be difficult to use the R-CNN model in real-time applications. Although the SSD model is faster than the others, since it is a light model using Mobilenet, the accuracy is low that it sometimes failed to detect a vehicle. YOLO is relatively slower than SSD, but it detects vehicles well without missing a car in every frame of video. The test results are shown in Figure 5. Comparing both mAP and FPS, we concluded that YOLOv4 is the most suitable model among the tested object detection models**.**

Table 2. FPS of Deep Learning Model

GPU : GeForce RTX 2080Ti

	YOLO v4	SSD	Faster-RCNN
FPS	82.1	105.14	36.32

Table 3. Evaluation of Deep Learning Models

Models	F1score	Precision	Recall	mAP
Yolo	0.96	0.93	0.98	98.19
SSD	0.88	0.90	0.87	90.56
Faster Rcnn	0.90	0.86	0.94	93.40

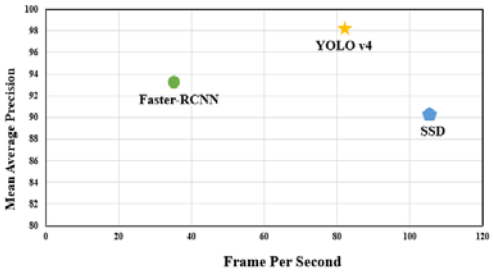


Fig 6. Performance of Deep Learning Model

4. Automobile License Plate Recognition Approach.

4.1 Data Set for Automobile License Plate Recognition (KETI-ALPR Dataset)

KETI-ALPR dataset contains information on license plates collected in South Korea. It consists of two main things. First, the position information of the license plate relative to the overall image is included. Second, information about the characters forming the license plate code is included. The license plate (LP) dataset consists of over 2000 images taken during the day and evening. Each image contains 1 to 4 license plates. Character dataset consists of over 3000 license plate images taken during the day and evening. Each image contains 7 to 9 characters.

Fig 7. Examples from KETI-ALPR dataset

4.2 Lincense Plate Detection

In this research, it is important to quickly and accurately detect relatively small license plates in

the image without setting an ROI (Region of Interest) on the input. Setting the ROI makes it easier to recognize license plates. However, if you set an arbitrary ROI, the system becomes limited, and detecting the vehicle area and setting it as ROI slows down the system.

We solved this by using YOLOv4. It has improved the detection of small objects compared to previous YOLO versions or SSD. Performance and speed can be tuned by adjusting the input network resolution. In our case, the size of 256x256 was selected because it was based on the performance of the NVIDIA Jetson TX2 board. If you are not using a resource-constrained platform, it is better to choose 512x512 or 608x608.

We were unable to conduct speed improvement tests because of camera device driver compatibility issues with TensorRT 4. However, it has been studied that using TensorRT 5 and YOLOv4 can provide speed improvements [9].

4.3 Character Recognition

At first glance, in a high-resolution image like 4K, the license plate looks clear. However, when cars are moving at high speeds, it is often difficult to recognize the license plates in its original size, as shown in **Fig. 8**. When using an SSD or previous version of YOLO, accuracy is poor because the features to be calculated are not clear. This can also be solved with YOLOv4. Similar to the LP detection setup, network resolution is set to 256x256 in order to speed up the processing time of the real-time system.

4.4 Experimental Results of License plate Recognition

We built the dynamic library by modifying only the necessary code of darknet and developed the system.

The license plate at the top of the image is small and the search rate tends to be down. Even if detected, the character part is crushed, making it difficult to recognize. The best recognition is when there is a license plate in the center or at the bottom as shown in **Fig. 8**.

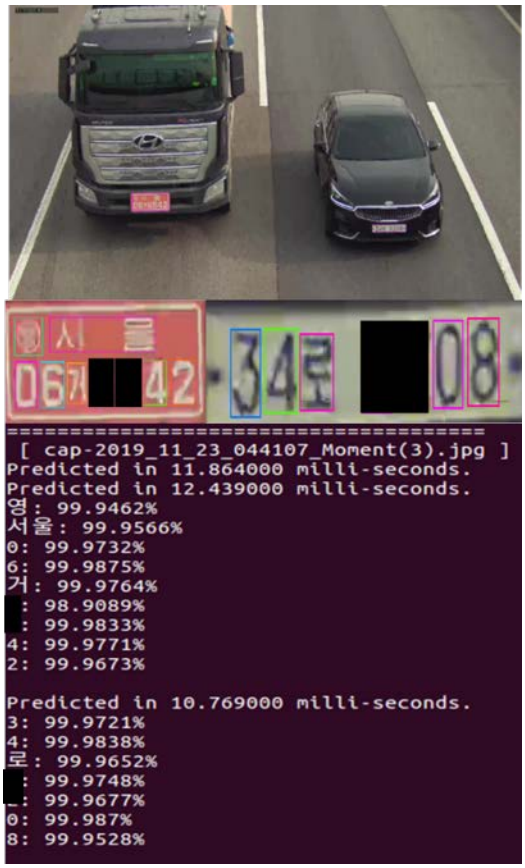


Fig 8. Results of KETI-ALPR system

The experimental results for accuracy and speed can be found in Tables 4 and 5. In general, the network resolution of 608x608 is more accurate than the case of 256x256, but in our experiment, 256x256 was better. This is probably because the input license plate images have widths ranging from 100 to 400 and heights ranging from 40 to 150, which is far smaller than the size of 608, so there are fewer features to calculate.

Table 4. The accuracy of character detector

Detector	F1-score	IoU	mAP@0.5
YOLOv3 (416)	0.93	84.91	71.53
YOLOv4 (256)	1	94.25	98.36
YOLOv4 (608)	1	89.55	98.12

Table 5. The processing time for each method

	GeForce RTX 2080 Ti		NVIDIA Jetson TX2	
Detector	Time (ms)	FPS	Time (ms)	FPS
SSD	51	19.6	74	13.5

SSD (TensorRT)	4.38	228	32.5	30.77
YOLOv4 (256)	9	111	105	9.5
YOLOv4 (608)	25	40	470	2.1

5. Conclusion

We tested several object detection methods to recognize the vehicle type and license plate. we compared its performance by processing speed and mean average precision. It can be seen that the recently released model, YOLO v4, has the best performance.

In YOLO version 4, features were predicted for each layer in the YOLO structure using FPN (Feature Pyramid Network). This solved the disadvantage of not catching small objects as high-resolution features are reflected in detection in YOLO. Also, version 4 was made for the purpose of being able to detect objects well with one GPU. In our experiment, we also measured the performance using one GPU, so it is judged that the learning of Yolov4 produced the optimal result.

We propose a fast and accurate automobile recognition system that is suitable for processing field data. By enabling small detection in the image, the burden of setting the ROI area was relieved, reducing the amount of computation and increasing the speed. In addition, the performance is improved by using a YOLOv4 detector that can recognize some contorted characters.

Acknowledgements

This work was supported by the IT R&D program of MOTIE/ KEIT. [10077018, Development of intergrated embedded platform and commercial prototype equipped with high-speed LPR function]

References

[1] Traore, Boukaye Boubacar, Bernard Kamsu-Foguem, and Fana Tangara. "Deep convolution neural network for image recognition." *Ecological Informatics* 48 (2018): 257-268.

[2] Girshick, Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. "Rich feature hierarchies for accurate object detection and semantic segmentation." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 580-587. 2014.

[3] Girshick, Ross. "Fast r-cnn." In *Proceedings of the IEEE international conference on computer vision*, pp. 1440-1448. 2015.

[4] Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster r-cnn: Towards real-time object detection with region proposal networks." In *Advances in neural information processing systems*, pp. 91-99. 2015.

[5] Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You only look once: Unified, real-time object detection." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779-788. 2016.

[6] Liu, Wei, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. "Ssd: Single shot multibox detector." In *European conference on computer vision*, pp. 21-37. Springer, Cham, 2016.

[7] Redmon, Joseph, and Ali Farhadi. "YOLO9000: better, faster, stronger." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7263-7271. 2017.

[8] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." *arXiv preprint arXiv:1804.02767* ,2018.

[9] Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao. "YOLOv4: Optimal Speed and Accuracy of Object Detection." *arXiv preprint arXiv:2004.10934*, 2020.

Self-attention Hough Voting Network for 3D Object Detection in Point Clouds

HyeOk Choi, SeHo Park, Kyung-Taek Lee
Korea Electronics Technology Institute, Contents Convergence Research Center
Seoul, South Korea
[e-mail: hochoi@keti.re.kr, sehopark@keti.re.kr, ktechlee@keti.re.kr]
*Corresponding author: Kyung-Taek Lee

Abstract

In this paper, we construct a network to perform an accurate 3D object detection task using self-attention module. The 2D detector based method, which is one of the 3D object detection methods that are widely used recently, shows good performance. However, in this method, if the 2D bounding box cannot be found in the scene, it becomes a critical factor for 3D object detection because the undetected 2D bounding box cannot be extended to the 3D bounding box. To solve this problem, 3D object detection method that uses a 3D point cloud directly as an input is used to avoid loss of information about the objects and configuration environments in the scene. To take advantage of the point cloud input, we propose a self-attention 3D object detection network using VoteNet[1] network and CBAM[2] which shows good performance in 3D object detection network and attention module. Our network was evaluated using SUN RGBD[4] dataset, achieving an accuracy of 58 mAP.

***Keywords*:** 3D Object Detection, Self-Attention, Deep Learning

1. Introduction

Object detection is a fundamental task that is widely used in computer vision. It detects the positions of all objects in a scene by creating a bounding box on an object with a category.

With the advancement of 3D sensors (e.g. Microsoft Kinect, Xtion) that provide color and depth information at the same time, it is easier to expand from 2D to 3D than a stereo sensor. It makes 3D object detection easier to detects the physical size and position of an object.

Recently, the emergence of CNNs (convolution neural networks) showed remarkable development in the field of computer vision.

However, CNN has limitations in its application due to the sparse 3D point cloud, so it is difficult to extract features from the scene. Various methods have been proposed to solve this problem.

[5, 6] propose a method of detecting an object using a 2d detector from an RGBD image. It increases the accuracy by utilizing the information of each pixel. However, this method has a limitation that if the 2D detector does not detect object with 2D bounding box, it cannot be extended to the 3D bounding box, which sacrifices accuracy.

On the other hand, [3, 1, 7] proposes a 3D nerual network to detect an object by receiving a 3D point cloud as an input.

This work was supported by the IT R&D program of MSIT/IITP. [R2020040040, Development of 5G-based 3D spatial scanning device technology for virtual space composition.

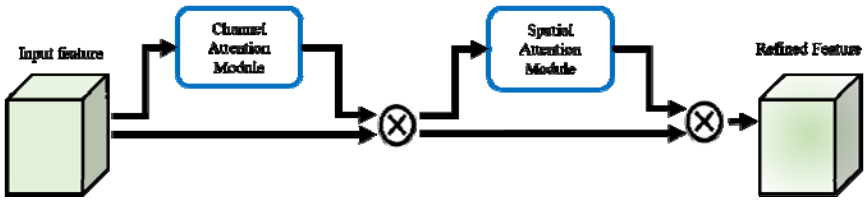


Fig. 1. Convolution Block Attention Module (CBAM)

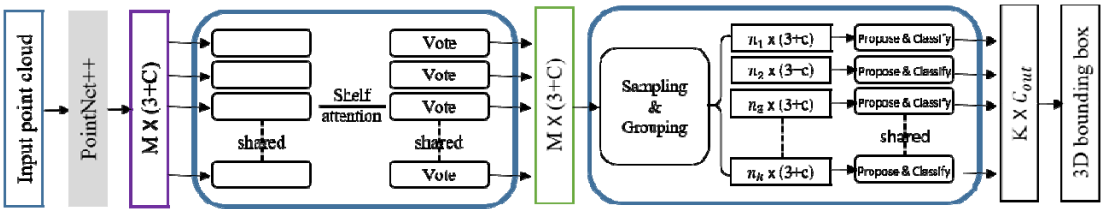


Fig. 2. Self-Attention Hough Voting Network

VoteNet[1] network, which has recently shown good 3D object detection performance, introduces a voting mechanism to solve the problem of poor accuracy due to the sparse point cloud characteristics. 3D point cloud feature extracted through PointNet++[3] is reduced the search space in the scene by using modified Hough Voting algorithm and improve detection accuracy through clustered votes.

In this paper, we propose a 3D object detection network using the attention module to increase the object detection accuracy of the baseline VoteNet[1] network. Proposed network helps in the proposal and classification of 3D objects by providing information about the set of point cloud that should be emphasized in the sparse point cloud of the scene.

Our network was evaluated on SUN RGBD[4] datasets. We used only the geometry information of the SUN RGBD dataset and showed better performance than the method [5, 8] which used RGB and geometry.

2. Method

2.1 VoteNet

In this paper, we use VoteNet[1] network, which shows good performance in the field of 3D object detection. It consists of a voting in point clouds part and an object proposal and classification from votes part.

Use PointNet++[3] as a backbone, which uses point cloud as direct input to extract point cloud features. By applying modified Hough Voting method, a new point close to the center of the object to be detected. These features can be used to generate votes through MLP(Multi-Layer Perceptron). Sampling and grouping these points, an object candidates are created. Finally, 3D objects are finally detected through proposal, classification, and regression.

2.2 CBAM

Fig. 1 shows Convolution Block Attention Module (CBAM[2]). CBAM proposes an effective attention module by sequentially configuring two modules, a channel attention module and a spatial attention module.

The relationship between feature channels is extracted through the channel attention module to perform computation on meaningful object information in the input scene. The spatial attention module is used to compute spatial information about where the object information is located.

Through the two-stage module, the features to be emphasized are trained more efficiently by increasing the performance of the representation computation of the CNN network.

2.3 Self-Attention Hough Voting Network

Fig. 2 shows proposed network architecture. The two-stage self-attention module is used in

	bathtub	bed	bookshelf	chair	desk	dresser	nightstand	sofa	table	toilet	mAP
VoteNet [1]	70.5	83.5	30.8	73.3	24.2	27.6	62.8	62.1	48.4	85.8	56.9
Ours	**76.2**	82.9	28.4	72.6	**24.8**	**34.1**	61.8	**63.8**	**48.9**	**86.9**	**58**

Table 1. 3D Object Detection evaluation results on SUN RGBD V1 validation set.

the voting module for voting on the input 3D point cloud. Through this end-to-end self-attention network architecture, meaningful 3D point clouds in the scene can be trained.

PointNet++[3] backbone extract global feature of 3D point clouds. This feature is used as input to the attention module to extract the emphasized 3D point cloud features. Voting on this feature helps to extract meaningful objects from the scene. Also, the voting mechanism can reduce the search space.

Clustered points through sampling and grouping are proposed and classified into 3D bounding boxes through multi-layer perceptron (MLP).

3. Experiment

We conducted the experiments using Pytorch running on a workstation with two GeForce RTX 2080Ti GPUs with CUDA 10.0, cuDNN 7.6.5.

3.1 Dataset

Our approach evaluate on SUN RGBD[4] datasets. It consists of 10,355 RGBD images with orientated 3D bounding boxes. It is a set of indoor scene composed of 37 classes. The evaluation is conducted for 10 common categories.

3.1 Results

Table 1 shows the performance evaluation results. As a result of training through the same workspace, our method achieve accuracy of 58mAP was better than 56.9mAP of VoteNet[1].

The feature of the attention module helped better understand the set of 3D point clouds of the scene by extracting features that need to be emphasizd in order to detect object.

4. Conclusions

In this paper, we proposed a 3D object detection method using VoteNet[1] that uses a point cloud as an input and attention module CBAM[2], which have recently shown good performance.

In order to solve the limitations of the sparse 3D point cloud, the attention module is used to extract and utilize features for objects that need to be emphasized in the scene.

Our method, confirmed the improved result with the accuracy of 58mAP by using only the geometry information of the 3D point cloud data.

Reference

[1] Ding, Zhipeng, Xu Han, and Marc Niethammer. "VoteNet: a deep learning label fusion method for multi-atlas segmentation." *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2019..

[2] Woo, Sanghyun, et al. "Cbam: Convolutional block attention module." *Proceedings of the European conference on computer vision (ECCV)*. 2018.

[3] Qi, Charles Ruizhongtai, et al. "Pointnet++: Deep hierarchical feature learning on point sets in a metric space." *Advances in neural information processing systems*. 2017.

[4] Song, Shuran, Samuel P. Lichtenberg, and Jianxiong Xiao. "Sun rgb-d: A rgb-d scene understanding benchmark suite." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.

[5] Song, Shuran, and Jianxiong Xiao. "Deep sliding shapes for amodal 3d object detection in rgb-d images." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.

[6] Deng, Zhuo, and Longin Jan Latecki. "Amodal detection of 3d objects: Inferring 3d bounding boxes from 2d ones in rgb-depth images." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017.

[7] Zhang, Zaiwei, et al. "H3DNet: 3D Object Detection Using Hybrid Geometric Primitives." *arXiv preprint arXiv:2006.05682* (2020).

[8] Lahoud, Jean, and Bernard Ghanem. "2d-driven 3d object detection in rgb-d images." *Proceedings of the IEEE International Conference on Computer Vision*. 2017.

Pose Tracking in Dynamic Scene using Monocular Camera and IMU sensor

Sukwoo Jung, Seho Park, and KyungTaek Lee
Contents Convergence Research Center, Korea Electronics Technology Institute,
Seoul, South Korea
[e-mail: ktechlee@keti.re.kr]
*Corresponding author: KyungTaek Lee

Abstract

Pose estimation of the sensor is important issue in many applications such as robotics, navigation, tracking, and Augmented Reality. This paper proposes visual-inertial integration system appropriate for dynamically moving condition. The orientation estimated from Inertial Measurement Unit (IMU) sensor is used to calculate the essential matrix based on the intrinsic parameters of the camera. Using the epipolar geometry, the outliers of the feature point matching are eliminated in the image sequences. The pose of the sensor can be obtained from the feature point matching. The use of IMU sensor can help initially eliminate erroneous point matches in the image of dynamic scene. The proposed procedure was implemented and tested, comparing with the existing methods. Experimental results have shown the effectiveness of the technique proposed in this paper.

***Keywords*:** pose tracking, pose estimation, IMU, sensor fusion

1. Introduction

The position and orientation information of the object is known as pose. Position represent the three dimensional location of the object, while orientation can be expressed as a series of consecutive rotations. Estimating the pose of sensor is important in navigation system. Pose estimation has been studied over the past few decades and various technologies have been researched. There are several approaches to estimate the pose of the sensor. The well-known method is using inertial or visual sensors.

Inertial Measurement Unit (IMU) sensors have been widely used for pose estimation in various applications including robots, aircrafts and navigation systems. IMU sensors normally contain three orthogonal accelerometer axis and three orthogonal gyroscope axis. IMU sensor can obtain the data in high-frequency rate and calculate the accurate results without complex operation. In addition, it can be small size, light weight, low cost, and adopt wireless communication technologies. Therefore, IMU sensors are able to track fast and abrupt movements. Furthermore, they are robust to the illumination changes and visual occlusions. However, it has a drift problem accumulated from the position estimation [1]. Therefore, additional sensor is normally used to aid the inertial sensor in the pose estimation. Some researches proposed to use Global Positioning System (GPS) [2] with inertial systems to overcome the problem of a long-term drift. However, the major problem of the GPS is intermittent loss of signal, especially for indoor environment [3].

Vision-based tracking has high accuracy for pose estimation [4-6]. It can precisely track the

This work was supported by the IT R&D program of MSIT/IITP. [R2020040040, Development of 5G-based 3D spatial scanning device technology for virtual space composition.]

relative motion between the camera and objects by measuring the movement of selected features in the consecutive image sequences. There are various methods in feature detection. Local detector such as SIFT (Scale-Invariant Feature Transform) [7] and SURF (Speeded-Up Robust Feature) [8] are invariant to the rotation and illumination [9]. BREIF [10], ORB [11], AKAZE [12] are also well known feature detectors for tracking. Those feature points can be used to calculate the pose of the camera [13]. However, vision sensor normally suffers from a lack of robustness, low frequency of data acquisition, and high computational cost. The main problem of vision-based tracking systems is instability with fast moving dynamic scene because of the loss of visual features. In addition, vision-based tracking systems are severely affected by occlusion matter. To overcome these problems, some researchers proposed approaches to take advantage of both the vision sensor and inertial sensor [14-16].

Typically, visual and inertial measurements are combined using a filtering framework. Kalman filter and its derivatives are normally selected to perform sensor fusion by integrating measurements from the various sensors. Kim and Park [17] developed a sensor fusion framework by combining Radar, Lidar and camera with Extended Kalman Filter (EKF) for pose estimation. Foxlin and Naimak [18] used both the inertial and vision data by a complementary Kalman filter for a pose estimation based on fiducials. Rehbinder and Ghosh [19] adopted gyroscope to measure the orientation, and the proposed vision-based system focused on correcting the drift of the inertial system. Although several researches showed that the inertial data can improve the vision-based tracking, the optimal way to integrate the data from inertial sensor and vision sensor is still unclear. Moreover, the integration of inertial data often causes instabilities in position estimation [20].

The purpose of this paper is to overcome the instabilities of sensor pose estimation by using both camera and IMU sensor. The proposed approach fused the data from the IMU sensor and the camera to estimate the pose of the sensor. As a result, the proposed method can estimate the pose of the sensor in real time and it can be applied for various applications including robot, autonomous driving, etc.

The remainder of this paper is organized as follows. Section 2 addresses the approach to the proposed pose estimation algorithm. Section 3 describes the methodology of the proposed method in details, and section 4 shows the results of the experimentation. Finally, the summary of this paper is given in the final section.

2. Approach to The Camera-IMU Data Fusion

2.1 Feature Matching

To calculate the pose of the camera from the image sequences, visual features should be obtained from the image sequences. Feature points can be detected by the detector including SIFT, SURF, AKAZE, etc. Since it has lower computational cost and moderate performance, ORB [11] is widely used in feature tracking applications. Then, descriptors are used to find the matching relations in the given sets of feature points. **Fig. 1** shows the example of feature point matching result with ORB detector and descriptor. Using feature point matching relations and camera calibrated parameters, the camera extrinsic parameters can be calculated.

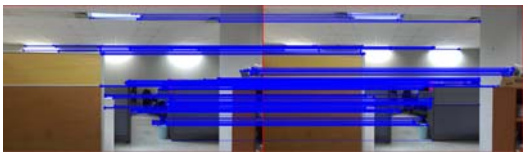


Fig. 1. The example of ORB feature matching result.

However, as mentioned earlier, the pose estimation results can be unstable because of the dynamic image sequences. Illumination change, blur, and motion of the object can adversely affect the position or matching relation of the feature points in the images. Consequently, the estimated pose results can be instable. To overcome this problem, we used epipolar geometry to remove the outliers of the feature points [13].

If the camera is not moving, the same feature points must lie on the same position in the next image frame. Then, the distances of the corresponding feature points are calculated and the feature points which have distance higher

than the threshold can be removed from the pose estimation. On the other hand, if the camera is moving, the same feature points must lie one the epipolar line in the next image frame. The feature points which have distance higher than the threshold are considered as an outlier. The epipolar line and the outliers are shown in **Fig. 2**.

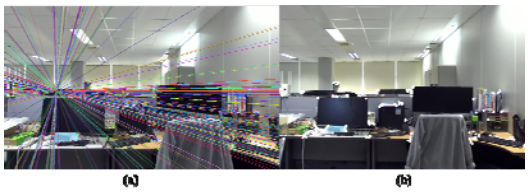


Fig. 2. The example of outlier removal process with epipolar geometry (a) epipolar line example (b) outlier detected by epipolar geometry. Red dots are considered as outlier.

However, if the feature points have a large portion of outliers in an image (e.g. blur, occlusion), It is difficult to determine the outliers properly with epipolar geometry. The epipolar line is drawn with the fundamental matrix which is calculated from the matching feature points and camera parameters. If the matching of feature points are not accurate, the fundamental matrix and epipolar line becomes inaccurate. As a result, the outlier rejection becomes challenging. To compensate the instability of the vision sensor, we decided to use IMU sensor as the initial estimator of the camera pose rather than feature matching.

2.2 Camera-IMU Fusion

Normally, IMU sensor has 6 DoF (degree of freedom) and can obtain linear and angular acceleration with high-frequency rate. As we aforementioned, inertial sensors offer good signals with high rate during fast motions but are sensitive to accumulated drift due to the double integration during the estimation of position. On the contrary, visual sensors provide precise ego-motion estimation with low rate in long term, but suffer from blurred features under fast and unpredicted motions. **Fig. 3** shows example of the pose estimated from the vision and inertial sensor. It shows the translation change in x-axis over frame. We obtained the data from each sensor with the same frame rate (30fps). As the sensor captured the image in static state, the ground truth must be zero. Both errors from each sensor are low because the camera is in static state, but they show distinctive different tendency. IMU data has high rate of flipping error and because of the drift error, the error is slightly increasing. On the other hand, vision sensor has no drift error, but sometimes error occur sharply because of the unpredicted motion of the image (e.g. illumination change, motion of the object). The aim of visual-inertial sensor integration is to overcome some fundamental limitations of visual-only tracking and IMU-only tracking using their complementary properties.

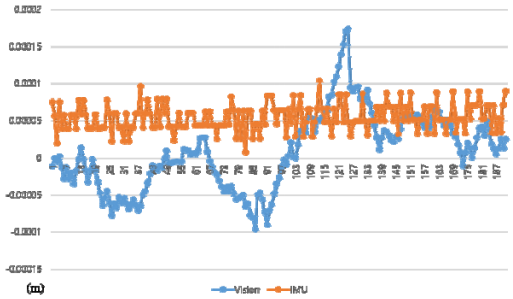


Fig. 3. The example of pose estimation from vision and IMU sensor.

As the IMU sensor can estimate the pose change of the consecutive image invariant to the abrupt motion, the transformation matrix obtained from IMU sensor can be used to the initial value for the fundamental matrix calculation. After removing the outliers with the initial fundamental matrix, the feature matching method will produce more precise transformation matrix. The overall procedure of the proposed method is shown in **Fig. 4**. Detail explanation will be given in the next section.

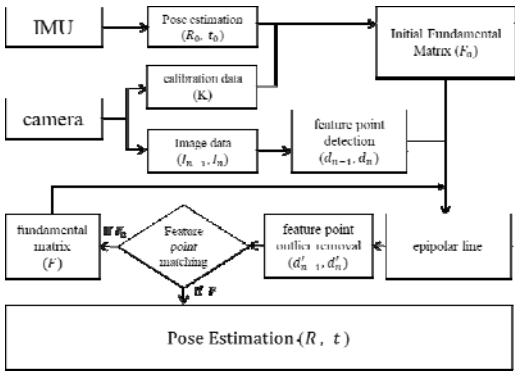


Fig. 4. Overall procedure of the proposed pose estimation method.

3. Proposed Pose Estimation Method

The IMU estimates initial value of rotation (R_0) and translation (t_0). Since the raw data has bias error, we applied Kalman filter [21] to reduce the error of the raw data. The initial fundamental matrix is calculated from the following equation.

$$F_0 = K^{-T} R_0 K^T [K R_0{}^T t_0]_X \quad (1)$$

where K are the calibration data of the camera. Calibration data is the intrinsic parameters of the camera. It can be calculated using bundle adjustment.

The feature points (d_{n-1}^l, d_n^l) are detected from the consecutive images (I_{n-1}, I_n) captured by the camera. The well-known detector ORB [11] is used in the proposed method. Using the initial fundamental matrix F_0 and the feature points d_{n-1}, the epipolar line can be drawn in the consecutive image I_n. Then, the epipolar line and feature point distance is calculated with the equation (2).

$$d = \frac{|au + bv + c|}{\sqrt{a^2 + b^2}} \quad (2)$$

where u and v are the coordinates of the feature point, and a, b and c represent the coefficient of the epipolar line, $ax + by + c = 0$.

Feature points with distance above threshold is considered as outlier. The optimal threshold value is obtained by experiments. We applied threshold value for 1.5 pixel, which means if the corresponding feature point is far from epipolar line more than 1.5 pixel, that feature point is considered as outlier in the pose estimation procedure.

After the outliers are removed from the feature points (d_{n-1}^r, d_n^r), these feature points matching relations are used to calculate the precise fundamental matrix (F). Then, the above procedure is repeated again to remove a few outliers remaining in the feature points. Finally, with the feature point matching relation, the pose of the sensor is estimated. In the next section, experimental results of the various pose estimation methods are demonstrated.

4. Experiments and Results

For the experiment, we used commercial sensor ZED 2 (StereoLabs, San Francisco, USA) which includes camera and IMU sensor in one platform. The raw sensory data are delivered to the host computer at 100 Hz via the USB 3.0 port. Camera can capture image sequences in 30fps with Full-HD resolution. We captured the both data in 30fps to synchronize the data from camera and IMU sensor. The time elapsed, which is the difference between the time instant when the acquisition process starts and the time instant when a new image frame is available, is collected together with the sensor data. The drift of the IMU sensor in translational axis is 0.35% and the drift in rotational axis is 0.005 °/m. The configuration of the experiment is shown in **Fig. 5**. After collecting data with the sensor, the data were transferred to the personal computer and the various pose estimation methods were tested with an i7-4790K processor with 16GB memory and a Windows 10 operating system.

Fig. 5. Configuration of the experiment.

The proposed method and the pose estimation method using only feature point matching (ORB) was implemented using OpenCV library [22]. Also, the previous method using Kalman filter to integrate the result of ORB and IMU data was implemented for comparison. We collected the data of moving sensor by 1m for 10 times. One set of data is shown in **Fig. 6**. The results of four pose estimation methods were compare with the given dataset. The position results of pose estimation methods must converge to 1m. The average positions of each algorithm are shown in **Table 1**.

As it can be seen in the graph, the pose estimation result from the IMU sensor has the largest error because of the drift. Pose estimation from the ORB feature matching has relatively precise result. However, sometimes the error fluctuates because of the instability of the image. Therefore, in a complex scene, the error of feature matching might increase significantly. As the two sensor data were used in the proposed method and Kalman filtering method, they

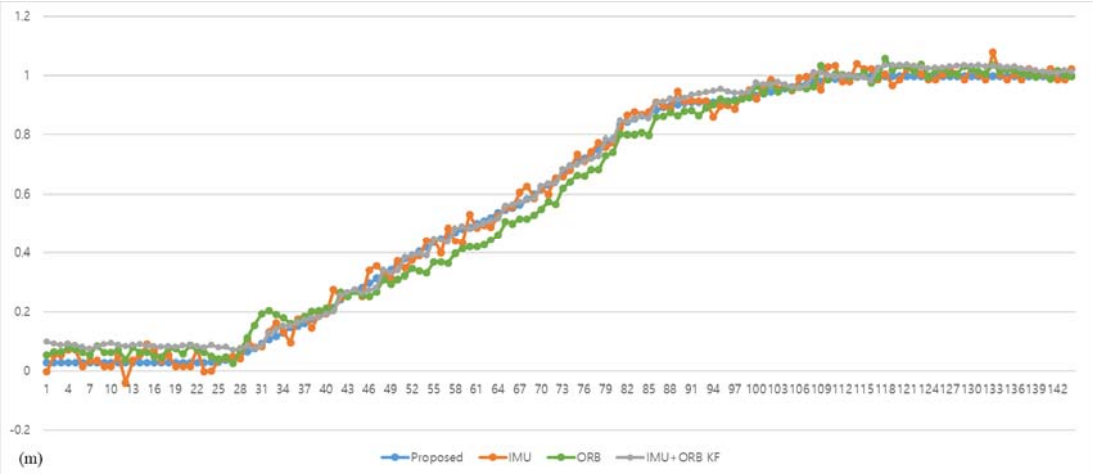


Fig. 6. Pose estimation result for the example dataset.

Table 1. Average position of the tested algorithms.

Algorithm	Proposed	IMU	ORB	IMU+ORB KF
Average position (m)	0.9996	1.0446	1.0085	1.0132
Standard Deviation	0.0004	0.0089	0.0251	0.0117

showed improved performance in the given experiment.

We also tested each algorithm with the data which were collected by the cart moving in a circle inside the office. The **Fig. 7** shows the position tracking result with the four algorithms including the proposed method.

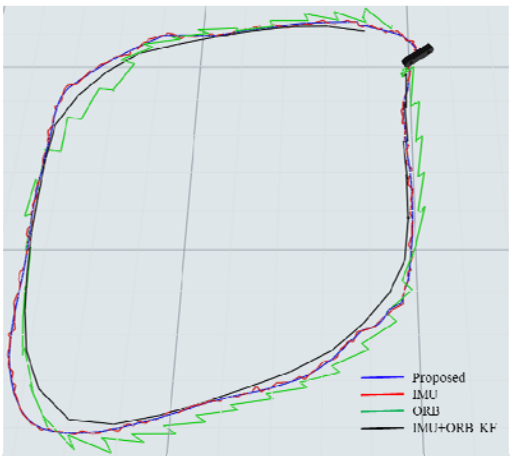


Fig. 7. Position tracking result of the indoor data.

As it can be seen in **Fig. 6-7** and **Table 1**, the proposed method shows the best performance.

5. Summary

In this Research, a new pose estimation method using image sequences and IMU sensor is proposed. The feature point matching method has a fundamental problem of instability because of the dynamic scene. The proposed method used IMU sensor to obtain the initial pose of the sensor. Then, the epipolar geometry is applied to eliminate the outliers in the feature points. With the feature points of outliers removed, matching relation generates the accurate pose of the sensor.

The proposed method used IMU sensor to compensate the pose estimation procedure, and it reduces the instability error in the pose estimation. The proposed method was implemented and was tested with the other methods. Experimentation shows that the proposed method provides better pose estimation results and is more robust compared to the other methods. Additional tests and research with the data from outdoor should be taken in the future work.

References

[1] A. R. Jiménez, F. Seco, J. C. Prieto, J. Guevara, "Indoor pedestrian navigation using an INS/EKF framework for yaw drift reduction and a foot-mounted IMU," *2010 7th Workshop on Positioning, Navigation and Communication*, pp. 135-143, 2010.

[2] F. Caron, E. Duflos, D. Pomorski, P. Vanheeghe, “GPS/IMU data fusion using multisensor

Kalman filtering: introduction of contextual aspects," *Information Fusion*, vol 7, no 2, pp. 221-230, 2006.

[3] F. Subhan, S. Ahmed, S. Haider, S. Saleem, A. Khan, S. Ahmed, M. Numan, "Hybrid Indoor Position Estimation using K-NN and MinMax," *KSII Transactions on Internet and Information Systems*, vol 13, no 9, pp. 4408-4428, 2019.

[4] M. Blösch, S. Weiss, D. Scaramuzza, R. Siegwart, "Vision based MAV navigation in unknown and unstructured environments," *2010 IEEE International Conference on Robotics and Automation*, pp. 21-28, 2010.

[5] M. E. Ragab, G. F. Elkabbany, "A Parallel Implementation of Multiple Non-overlapping Cameras for Robot Pose Estimation," *KSII Transactions on Internet and Information Systems*, vol 8, no 11, pp. 4103-4117, 2014.

[6] L. Li, Y. Liu, T. Jiang, K. Wang, M. Fang, "Adaptive Trajectory Tracking of Nonholonomic Mobile Robots Using Vision-Based Position and Velocity Estimation," in *IEEE Transactions on Cybernetics*, vol 48, no 2, pp. 571-582, 2018.

[7] D. G. Lowe, "Object recognition from local scale-invariant features," *Proc. of the Seventh IEEE International Conference on Computer Vision*, vol 2, pp. 1150-1157, 1999.

[8] H. Bay, A. Ess, T. Tuytelaars, L. V. Gool, "SURF: Speeded Up Robust Features," *Computer Vision and Image Understanding*, vol 110, no 3, pp. 346-359, 2008.

[9] S. Jung, S. Song, M. Chang, S. Park, "Range Image Registration based on 2D Synthetic Images," *Computer-Aided Design*, vol 94, pp.16-27, 2018.

[10] M. Calonder, V. Lepetit, C. Strecha, P. Fua, "BRIEF: Binary Robust Independent Elementary Features," *ECCV 2010: Proc. of the 11th European Conference on Computer Vision*, pp.778-792, 2010.

[11] E. Rublee, V. Rabaud, K. Konolige, G. Bradski, "ORB: An efficient alternative to SIFT or SURF," *2011 International Conference on Computer Vision*, pp. 2564-2571, 2011.

[12] P. Alcantarilla, J. Nuevo, A. Bartoli, "Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces," *in Proc. British Machine Vision Conference 2013*, vol 13, pp. 1-11, 2013.

[13] S. Jung, Y. Cho, D. Kim, M. Chang, "Moving Object Detection from Moving Camera Image Sequences Using an Inertial Measurement Unit Sensor," *Applied Sciences*, vol 10, pp. 268, 2020.

[14] J. L., J. A. Besada, A. M. Bernardos, P. Tarrío, J. R. Casar, "A novel system for object pose estimation using fused vision and inertial data," *Information Fusion*, vol 33, pp. 15-28, 2017.

[15] Y. Tian, Jie Zhang, J. Tan, "Adaptive-frame-rate monocular vision and IMU fusion for robust indoor positioning," *2013 IEEE International Conference on Robotics and Automation*, pp. 2257-2262, 2013.

[16] K. Kumar, A. Varghese, P. K. Reddy, N. Narendra, P. Swamy, M. G. Chandra, P. Balamuralidhar, "An Improved Tracking using IMU and Vision Fusion for Mobile Augmented Reality Applications," International Journal of Multimedia and its Applications, vol 6, num 5, 2014.

[17] T. Kim, T.H. Park, "Extended Kalman Filter (EKF) Design for Vehicle Position Tracking Using Reliability Function of Radar and Lidar," *Sensors*, vol 20, pp. 4126, 2020.

[18] E. Foxlin, L. Naimak, "VIS-Tracker: A Wearable Vision-inertial Self-Tracker," *Proc. of IEEE VR2003*, vol 1, pp. 199, 2003.

[19] H. Rehbinder, B. K. Ghosh, "Pose estimation using line-based dynamic vision and inertial sensors," in *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 186-199, 2003.

[20] G. Qian, R. Chellappa, Qinfen Zheng, "Bayesian structure from motion using inertial information," *Proceedings. International Conference on Image Processing*, 2002.

[21] F. M. Mirzaei, S. I. Roumeliotis, "A Kalman Filter-Based Algorithm for IMU-Camera Calibration: Observability Analysis and Performance Evaluation," in *IEEE Transactions on Robotics*, vol 24, no 5, pp. 1143-1156, 2008.

[22] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

Automatic Video Splitting System based on Inter-frame Similarities

Jaewon Lee, Dalwon Jang and JongSeol Lee
Communications and Media R&D Division, Korea Electronics Technology Institute
Seoul, South Korea
[e-mail: jwlee0121@keti.re.kr, dalwon@keti.re.kr, leejs@keti.re.kr]
*Corresponding author: Jaewon Lee

Abstract

Recently, with the development of video content market and related technologies such as editing, producing personal videos, the importance of the pre-processing in the video creating industry is being emphasized. This paper deals with a system that divides video through key-frame extraction and provides it to users. For this, the video is first divided into frames, and the similarities between adjacent frames is calculated through a state-of-the-art pre-trained model, DISTS, which utilizes convolutional neural network and a color space histogram. As mentioned, the inter-frame similarity calculation goes through a total of two-steps, and the key-frame extraction and division of video for each section are performed through simple post-processing that removes meaningless frames. We implemented the proposed video splitting system mentioned above using VSUMM dataset.

***Keywords*:** key-frame extraction, video split, shot boundary detection, image similarity

1. Introduction

Nowadays, as creators who produce and share various contents increas, the need for a technology for reprocessing video contents is on the rise. In particular, prior to producing video content in order to reprocess the video, it is necessary to preemptively divide the video into shot units. There have been many studies on extracting key-frames or summarizing video through shot boundary detection[1-4], but few are provided to users in the form of splitted video.

In this paper, we will introduce a system that automatically divides video by extracting key-frames from video based on inter-frame similarities. The inter-frame similarities is calculated by sequentially applying the algorithm introduced in [5], which considers the structural and textural similarity of the two images, and the analysis method through the histogram difference in the color space.

The composition of this paper is as follows: The proposed system overviews and details are described in Section 2. The verification of the system is given in Section 3. Lastly, Section 4 concludes this paper.

2. System overview

2.1 Inter-frame similarities

It is important to extract features of images for frame-by-frame analysis. There are various methods such as edge investigation, calculating and comparing optical flow for feature extraction, but recently, Convolutional Neural Network(CNN) have been widely used. Algorithms using CNN are performing better in various fields than conventional methods. Therefore, we used the algorithm proposed in [5], a method that compares the similarity of two images with weights pre-trained from VGGNet. In addition, through the RGB characteristics of

This research is supported by Ministry of Culture, Sports and Tourism and Korea Creative Content Agency (Project Number: R2020040068)

the color image, two-step key-frame extraction was performed using the difference in the RGB histogram of the adjacent frames with the primarily extracted candidate frames.

2.2 Extracting key-frames

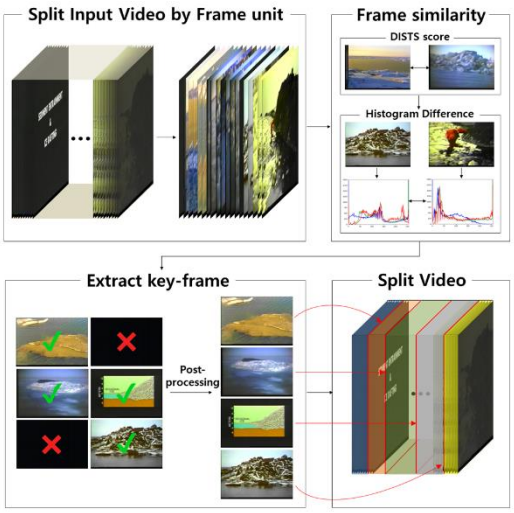


Fig. 1. Structure of the Video splitting system

In **Fig. 1**, the structure of the automatic video splitting system is shown. Before calculating the similarities between frames, the input video is divided into each image frame. Next, for all frames, the similarity between the target frame and the next frame is calculated by the method proposed in [5], and frames exceeding the specific threshold θ_1 are selected. Secondly, through RGB histogram analysis, final key-frames are selected with threshold θ_2. The equation of histogram difference is expressed by

$$d(f_n, f_{n+1}) = \sqrt{\sum_{i \in R,G,B}^{M} [h_i(f_n) - h_i(f_{n+1})]^2} \quad (1)$$

where i, n and M denote the color space of RGB, the frame index and the number of bins respectively. After the above process is over, meaningless frames such as black screens were excluded from key-frames by post-processing. **Table 1** is a pseudocode showing the entire process of the proposed system. As shown in Table. 1, if m key-frames are extracted from **S**, the number of divided videos generated from the corresponding system becomes m+1.

Table. 1. The Pseudocode of the proposed system

Algorithm

Input: the whole video file.
Output: Splitted video result.
Set **S** to store primarily extracted frames
while (video frames) **do**:
  Calculates $DISTS(x(n), x(n+1))$
    if $DISTS(x(n), x(n+1)) > \theta_1$:
      Calculates $d_{RGB}(x(n), x(n+1))$
        if $d_{RGB}(x(n), x(n+1)) > \theta_2$:
          Select and save n^{th} frame to **S**
        end if
    end if
end while
k **for** length(**S**)**:**
  Split video with index from **S**[k] to **S**[k+1]
  Save video in the form of .mp4

3. Experimental Results

Fig. 2. Test results on VSUMM dataset. (a) DT, (b) OV, (c) STIMO, (d) FFMPEG, (e) Proposed

We verified our video splitting system with VSUMM dataset[6]. VSUMM consists of 50 videos, all of which are 3 minutes or less long and 352x240 pixels, 30fps in MPEG-1 format. Some of them were tested, and the sample test results are shown in **Fig. 2**. As the parameters in the proposed system, 0.25 and 55,000 were used for θ_1 and θ_2. In the case of RGB histogram analysis in eq. (2), the size of bins, M is used with 256 each. User-selected key-frames as GroundTruth were taken from the VSUMM dataset.

To verify the proposed algorithm, we compared it with other methos including DT[7], OV[8], STIMO[9], FFMPEG[10], and the result of the test is shown in **Fig. 2**. As a result of testing the proposed algorithm, it was confirmed that the key-frames extracted by the proposed method was closer to the key-frames selected by the user than the results of (a)~(d). Most of the results showed that there was a missing frame in the entire content of the video or tended to be focused on specific shot.

5. Conclusions

In this paper, we proposed an automatic video splitting system that provides divided videos by extracting key-frames through two-stage inter-frame similarities. In particular, we proposed a method of calculating the similarity between images using CNN, which is the latest technique for feature extraction, and analyzing histograms in RGB color space. The test results with the proposed method showed better results than other methods in splitting the videos.

References

[1] C. Gianluigi and S. Raimondo, "An Innovative Algorithm for Key Frame Extraction in Video Summarization", *Journal of Real-Time Image Processing*, 1(1), 69-88, 2006.

[2] Y. Huo, Y. Wang, and H. Hu, "Effective Algorithms For Video Shot And Scene Boundaries Detection", *2016 IEEE/ACIS 15th International Conference on Computer and Information Science(ICIS)*, pp. 1-6, 2016.

[3] H. Wei, B. Ni, Y. Yan, H. Yu, X. Yang and C. Yao, "Video Summarization via Semantic Attended Networks", in *Proc. AAAI*, pp.216-223, 2018.

[4] C. Huang and H. Wang, "A Novel Key-Frames Selection Framework for Comprehensive Video Summarization", *IEEE Transactions on Circuits and Systems for Video Technology*, 30.2, pp.577-589, 2019.

[5] K. Ding, K. Ma, S. Wang, and E. P. Simoncelli, "Image Quality Assessment: Unifying Structure and Texture Similarity", *arXiv preprint arXiv:2004.07728*, 2020.

[6] [Online] https://sites.google.com/site/vsummsite

[7] P. Mundur, Y. Rao, and Y. Yesha, "Keyframe-based video summarization using Delaunay clustering," *Int. J. Digit. Libraries*, vol.6, no.2, pp.219–232, Apr. 2006.

[8] [Online] https://openvideo.org

[9] M. Furini, F. Geraci, M. Montangero, and M. Pellegrini, "STIMO: STIll and MOving video storyboard for the Web scenario," *Multimedia Tools Appl.*, vol.46, no.1, pp.47–69, Jan. 2010.

[10] [Online] https://ffmpeg.org

Demonstration of High-precision 3D Spatial Data Sharing Platform Services

Ilryong Kweon, Minho Son, Jungsun Ryu
Strategy dept. PODO inc.
Seoul, South Korea
[e-mail: mhson@ipodo.co.kr, irk@ipodo.co.kr,jsyoo@ipodo.co.kr]
*Corresponding author: Minho Son

Abstract

This study is to transform high-precision three-dimensional spatial information for virtual training system and to develop land and flight mobile virtual training system based on high-precision three-dimensional spatial information. Development of technology for converting and utilizing high-precision 3D spatial data for virtual training and collecting high-precision 3D spatial data and information on conversion requests using virtual training systems to generate spatial data interlocking virtual training data and study the technology development and utilization system that collects, transforms, transforms, visualizes and distributes them.

***Keywords*:** High precision 3D, cloud platform, virtual training, GIS

1. Introduction

Recently, the demand and need for high-precision 3D spatial data have increased in society. In order to meet these social needs, this study seeks to study the following elemental technologies: Expected rapid spread of spatial data, expected increase in information efficiency, and reduced social costs through user-centered distribution of high-precision 3D spatial data contents through the platform. Building virtual flight training contents, saving more than 70% of time and securing competitive edge in related industries by utilizing national spatial data. Reducing the cost of building virtual training background data is expected to contribute to the development of the industry by facilitating the operation and support of real-time training systems for workers, which lead to the improvement of productivity and efficiency, leading to the development of the industrial economy. Korea is expected to achieve 10-15% GDP growth by introducing extensive virtual training in industry and manufacturing facilities. Using virtual training services, it is expected that industries will be promoted by creating new jobs and upgrading industrial technologies through the pattern of industry-friendly manpower required by major industries such as manufacturing and root industries. Realistic and unmanned virtual training to prevent accidents and reduce social accident costs. Effect of practical policy support that conforms to the government's policy stance, such as building and opening up high-quality spatial data, revitalizing spatial data convergence industries, fostering virtual reality industries, and responding to scientific and strategic disasters. It faithfully performs its role as a core infrastructure for the creative economy by creating an ecosystem for the spatial information industry and creating new industries. Based on data sharing services using advanced spatial data simulation, it is expected that various markets related to spatial information will be created, small and medium

This work is supported by the Korea Agency for Infrastructure Technology Advancement(KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 20DRMS-B147026-03

enterprises will be fostered, and jobs will be created through convergence services with other information owned by existing industries. Maximize support for individual and small spatial data providers and enhance the public's awareness of the spatial information industry. Securing the basic status of the spatial information industry, strengthening the profitability of the spatial information industry as a whole, and securing a stable platform for providing services. Select and link platforms that are considered to be in high demand for services, and provide customized spatial information contents to the private sector to meet social needs. Prevent duplicate investment in the establishment of high-precision realistic contents by each ministry by sharing and utilizing national spatial data. Activate the industry and improve the image of the national brand by using various contents using national spatial data.

2. Related research

The state has tried to provide spatial data services in various ways, but with a scattered service system, Due to the lack of up-to-date and reliable information, it is difficult to utilize spatial data.

The country by integrating services scattered here and therein.Public and private public information in one place, at a time, anyone. A national spatial information portal was established for easy utilization. From January 1, 2016, the open distribution channel of spatial data has been unified into the national spatial data portal, and the distribution of spatial information open datasets, inquiry of spatial data list, and checking spatial data.Data rooms such as standards and statistics, maps (baro e-map, satellite images, national territory information basic map, flight safety zone, LOD service), open APIs, and numerical control types are also composed of searches by map. Space information is an important resource that enables intelligent services throughout society, including roads, automobiles, and cities, in the era of the fourth industrial revolution. The National Spatial Information Center continuously collects 2,560 kinds of spatial data held by 160 institutions, including central government agencies, public institutions, and local governments, to promote the private use of such spatial information.

It is pushing to expand its opening through the portal.

The purpose of this research is to develop a technology development and utilization system that collects high-precision 3D spatial data and generates spatial data-linked virtual training data by collecting conversion request information using virtual training system, and collects, transforms, transforms, visualizes and distributes them.

3. Data collection system

The establishment of storage space for the storage and processing of collected data defines the language of spatial big data basic operators and spatial big data query through spatial big data query processing engine.

Based on this, language parsing and basic search functions of spatial big data query were developed (area, real-time data, video, topography, buildings, and facilities data). The development of server interfaces for collecting and storing original data for virtual training is carried out through spatial data service data definition and spatial data collection processing engine to collect data from other ministries

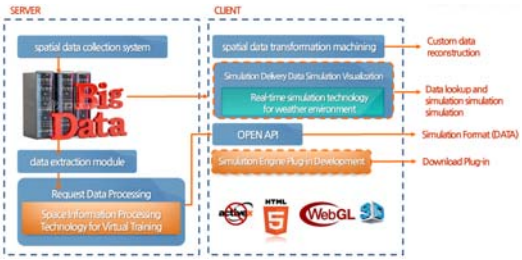


Fig. 1. High-precision 3D spatial information platform configuration for virtual training

Fig. 2. High-precision 3D spatial information platform configuration for virtual training

4. Data processing and management

In the area of visualizing spatial information for virtual training, the technology for visualizing spatial information for virtual training was developed. The spatial data inquiry function allows users to inquire and manage data according to the classification and layer designation of user data. Customized simulation environment is provided to users through the transformation of buildings and facilities for utilizing virtual training.

Library-based building random addition/deletion/change function is provided for building change function[**Fig. 3**].

Since data provided to users may need to be transformed according to physical and environmental conditions, it provides the ability to transform space data for virtual training to provide data tailored to that situation. To this end, this study provides features for editing topographic data (flatting, overwriting terrain, and correcting errors) [**Fig. 4**].

Fig. 3. Building editing

5. Data processing and management

In this study, For visualizing three-dimensional model data, the three-dimensional model data was designed to have a defined volume and shape in three-dimensional space and divided into ground and underground location objects based on the ground. The three-dimensional model has a three-dimensional space area according to the specified model and is used for determining proximity to the training area based on the space area.

Fig. 4. Terrain change editing

Data Delivery for Virtual Training Open API Development provides design and development of Open API operation server architecture and simulation data requests and transforms [Figure 5].

To this end, simulation data was defined and simulated data retrieval requirements (interface). Other simulation data API definitions (attribute lookup, area lookup, file download, etc.); user download data generation module development; response code definition (normal/above), Open API operation server construction; administrator page development for Open API server operation; web page-based server environment setting; design and development for processing Open API service connection statistics (logging call status for each function; log and log and statistical data processing DB supporting server operation and statistical data processing).

When the user selects an area to download after searching or moving to a desired location, the user can visually download and download the

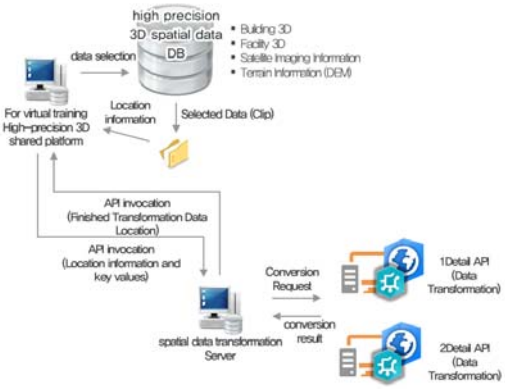


Fig. 5. API service system(structure)

data of the intersection with the set buffer on the web screen. In this system, visualized modeling data [**Fig. 6**] and download results were implemented using line selection as line input.

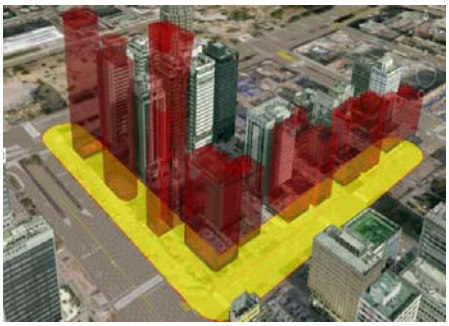


Fig. 6. Data extraction visualization

6. Conclusions

The following technical effects are expected from this study: Improving the quality of content and securing core technologies for high-performance, high-function virtual training simulators by securing technology to create virtual training contents based on high-precision 3D spatial data possessed by the state. When utilizing national high-precision 3D spatial data, the quality of virtual training contents is expected to be improved by 30% and the time cost of deployment is expected to be reduced by 70% compared to the content it builds on its own. Securing key technologies for realistic video processing based on spatial data that can express virtual information, contents, and interfaces in spatial data in a realistic manner. 3D high-precision spatial data can be retrieved and provided through a single window without having to look at all scattered data production agencies for the purpose of searching the underlying data for virtual training content. High-precision 3D spatial data is provided through data visualization in the aspect of simulation demanders, providing data retrieval function so that the results of data utilization can be predictable. Data reconstruction function allows various changes based on reality to be applied so that existing spatial data can be re-configured for the purpose of virtual training contents. Based on high-precision 3D spatial data based on actual measurement, virtual training contents are organized to provide realism and realism of simulation. Increased utilization of high-precision spatial data in simulation fields such as virtual training is a technical demand for the development of high-precision 3D spatial data update and upgrading technology. Securing infrastructure for producing high-precision 3D spatial data-based spatial cultural contents and eset store services to foster the industry. Reducing the cost of producing virtual training content lowers the entry barrier to the virtual training S/W market, providing an opportunity for small and medium-sized enterprises to actively participate in the market. Improving the quality of virtual training contents leads to the improvement of virtual training simulation S/W quality.

References

[1] Lee Byung-gil. "Final report on technology based on high precision 3D construction" Ministry of Land, Infrastructure and Transport, 2011.

[2] Kang Sang-gu, “The Final Report on the New Value Creation Foundation of Spatial Information by High Precision Positioning,” Ministry of Land, Infrastructure and Transport, 2016.

[3] Molenaar, M., “A formal data structure for 3D vector maps.” Proceedings of EGIS’90, Vol. 2, 1990.

[4] Pilouk, M. "Integrated modeling for 3D GIS." Ph.D. Dissertation, 1996.

[5] Zlatanova, S .. "3D GIS for urban development." Ph.D. Dissertation. 2000.

[6] Coors, V.. "3D GIS in Networking environments." Computer, Environment and Urban Systems, Vol. 27, 2003.

KSII The 12th International Conference on Internet (ICONI) 2020.

Semantic REST API: Towards Interoperability in Semantic and Non-semantic Environments

Sunil Kumar, Jieun Lee and SeungMyeong Jeong
Autonomous IoT Research Center, Convergence System R&D Division,
Korea Electronics Technology Institute
Seoul, South Korea
[e-mail: sunil75umar@keti.re.kr, je.lee@keti.re.kr, sm.jeong@keti.re.kr]
*Corresponding author: Sunil Kumar

Abstract

With the emergence of semantic web technologies, different tools, and database management systems provide APIs to access the semantic data. However, accessibility is often limited to a specific domain and standard-based environments. In addition, the utilization of those APIs require complete knowledge of semantic web and SPARQL, making it infeasible to users for seamless interworking with semantic data. This paper proposes a Semantic REST API, as a solution to improve accessibility between semantic and non-semantic environments. To achieve easier interpretation, six basic functionalities along with their respective resource description has been defined for the retrieval and traversal of semantic data. As proof of concept, a prototype has been developed, demonstrating the applicability of these six functionalities.

***Keywords*:** Semantic data access, Semantic REST API

1. Introduction

Accessing semantic data through SPARQL has been a convention for long time. And many triple stores provide REST APIs as a web service. However, accessing the data requires the client to have SPARQL query embedded as a request parameter. For a client with a non-semantic environment, different complexities become crucial to handle just to access the data such as, familiarity with RDF, OWL, Literals, graph structure and traversals etc., [1] which then requires additional tools and components for query formulation and data interpretation.

The main goal of Semantic REST API is to bridge the gap between semantic and non-semantic environments. In addition, it can support in modularization, minizing the interdependence of non-semantic client applications in terms of data interpretation, query translation, etc. This can be realized in **Fig. 1**, where client-side in 1(b) is simplified as compared to 1(a).

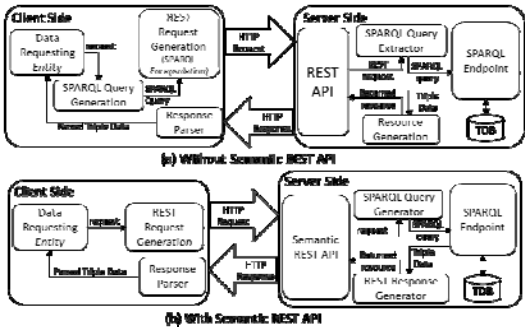


Fig. 1. Client and server interaction

2. API Description

2.1 Resource Description

In order to represent the semantic data in terms of REST resource structures, different resources have been defined. **Fig. 2** visualizes the semantic data both in terms of stored in triples database (TDB) and defined as REST resources. Here the term "Entity" refers to the data (triples) describing a particular Class, Named Individual, Object Property or a Datatype Property. **Fig. 2(a)** shows 7 main types of resources based on the sematic data stored in TDB. Among them, "Prefixes" and "Entity Info" have sub-types: simple and normal, in order to show the different possible formats, in which a resource can be defined. Normal format contains explicit definitions of keys to provide comprehensive description of data. Whereas simple format excludes those keys to support brief resource structure.

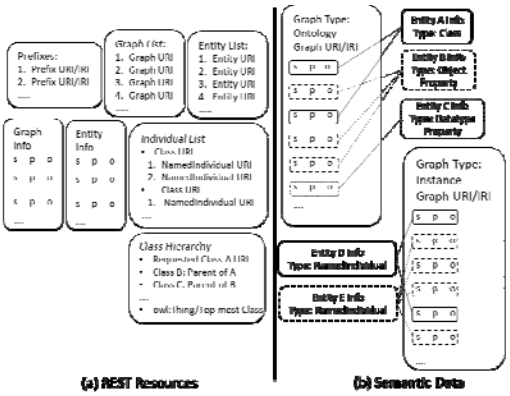


Fig. 2. Semantic data representation. Where, "s p o" refers to a single triple.

2.2 API Functions.

Based on the resources, six basic API functions have been defined: Graph List, Entity List, Graph Info, Entity Info, Individual List and Class Hierarchy. Although, there can be other functions defined such as graph traversals, inquiring complex assertions etc., we stick to the basic functionality to limit the scope of this work. Here all the functions are mapped to GET requests as in this work we have only focused on information retrieval.

The "Graph List" function retrieves the list of all the Graph URIs in the TDB based on the given graph type and the list of search keywords. Generally, there are two graph types: ones which store the ontology/vocabulary and the ones which store the assertions based on those ontologies. The "Entity List" function retrieves the list of entity URIs based on the given entity type and keywords. Here the type can be of the following: Class, Object Property, Datatype Property and Named Individuals. The "Graph Info" function returns all the triples of a particuar graph, based on the given graph URI. Similarly "Entity Info" function returns all the triples describing an entity, based on the given entity URI. The description will involve the entity type and the other related assertions. This can be considered similar to the DESCRIBE SPARQL query. The "Individual List" function retrieves all the URIs of Named Individuals, based on the given Class URIs. Finally, the "Class Hierarchy" function retrieves the list of the Class URIs, starting from the given Class URI till the top most class in the ontology. Note that this function does not include the sibling Classes. The returned list only shows parent child relationship.

Each response of the API function also include the prefix list, so the response is concisely formated using prefixes instead of complete namespace URIs. In addition, each request can include a limit as a query parameter, which limits data values in the response.

3. Proof of Concept

In order to articulate the applicability of the API functions, a prototype has been developed, using java spring as application framework and apache tomcat as servlet container. The triple database deployed is virtuoso (open source version: 7.2). For each API function, SPARQL queries have been defined to retrieve the information from database, taking into account the query parameters orchestrating these queries. The triple store contains the data, which is annotated from IoT sensor data acquired from parking lot in Yatap, Seongnam-si, Gyeonggi-do Korea. The ontologies used for annotating the data are discussed in a work by J. An et al. [2]. Among them, two ontologies are used here. One is common ontology, which defines all the top level generic concepts, and the second is parking lot ontology, which is extended from common ontology to define concepts related to parking domain. This prototype has been adopted as an API in web system for smart city data [3].

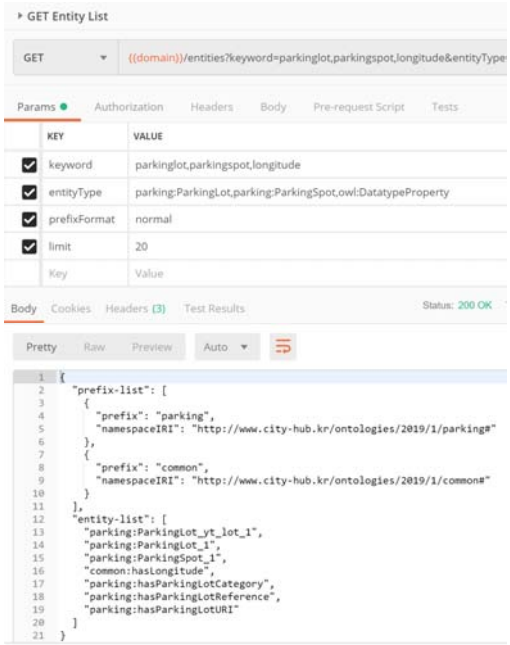


Fig. 3. API GET request to retrieve the entity list.

Fig. 3 shows the postman request for the API function: Entity List. Here the second query parameter contains the list of entity types, where the first two are the URIs of classes from parking ontology. In the in the "entity-list" array, the fourth element is of type owl:DatatypeProperty (third element in "entityType") and also matches with the given keyword "longitude" in the query variable. The "prefixFormat" query variable is "normal". The other possible value is "simple", where the list element excludes the keys "prefix" and "namespaceIRI". In that case, the value of the "prefix" becomes the key and value of "namespaceIRI" becomes the value of that key.

Considering this request, one question arises that how the prefixes can be managed, because SPARQL response returns complete URIs and prefix information is not stored as the part of graph data. There are two possible solutions for managing the prefixes. One is traversing the URIs in the response to provide an arbitrary prefix (for e.g. prefix1, prefix2, etc.) for each namespace URI. Second is to provide support to locally store the prefixes. In case of first solution, no namespace URI will be left without a prefix, however, it may not be user friendly, as it may require more processing when there are many prefixes involved. The second solution can be user friendly, as those prefixes might be easily recognizable, however, some namespace URIs may be left without a prefix, hence showing complete URI in the response.

Another featuer, which can be added is the type-wise organizing the response. So in case of **Fig. 3**, the entity list can be sub-divided based on "entityType". This will ease the process of parsing the information from user perspective, however, it may cause a delay in sorting the response, if the number of entity types increases.

4. Conclusions

This paper proposes Semantic REST API, to provide an alternate accessibility solution to formulating SPARQL queries, for the users, especially non-semantic standard based environments. In order to represent semantic data in REST environment, seven resources has defined, and their retrieval is supported by defining six basic API functions. The support for complex SPARQL queries, such as retrieving complex graph traversals and assertions, as well and INSERT and UPDATE functions are considered for future extension of this work.

ACKNOWLEDGMENT

This work is supported by Smart City R&D project of the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (MOLIT), Ministry of Science and ICT (MSIT) (Grant 18NSPS-B149388-01).

References

[1] Schröder, Markus, et al. "Simplified SPARQL REST API." *European Semantic Web Conference. Springer, Cham*, 2018.

[2] J. An, S. Kumar, J. Lee, S. Jeong and J. Song, "Synapse : Towards Linked Data for Smart Cities using a Semantic Annotation Framework," *2020 IEEE 6th World Forum on Internet of Things (WF-IoT), New Orleans, LA, USA*, 2020, pp. 1-6.

[3] J. Lee, J. An, S. Jeong and J. Song, "i-LOD: Industrial Linked Open Data System for Semantic Integration of Industrial Real-time Data in Smart City," *2019 IEEE International Conference on Industrial Internet (ICII), Orlando, FL, USA*, 2019, pp. 297-298.

Analysis of the moderating effect of fields, grades, and experience on the platform quality and reuse intention of computational science and engineering platform

Nam-Gyu Kim, Noo Ri On, Myoung-Ju Koh, Inho Jeon, and Jong Suk Ruth Lee
Center for Computational Science Platform, Korea Institute of Science and Technology Information
245 Daehak-ro, Yuseong-gu, Daejeon 34141, South Korea
[e-mail: {ssgyu, claireon, myju, inojeon, jsruthlee}@kisti.re.kr]
*Corresponding author: Jong Suk Ruth Lee

Abstract

The EDISON platform was developed and has been in service since 2011. It is the only computational science and engineering platform in Korea that is being used by a total of 79,000 cumulative users. Most users on the EDISON platform are science and engineering university students who use the platform for educational and research purposes. Surveys on these users have been conducted for many years, but analysis studies from various viewpoints are insufficient. In this study, we analyze how the background of users affect their intention to reuse the platform. To perform the analysis, a survey was conducted on the participants of the EDISON SW Utilization Contest over three years from 2018 to 2020. User's intention to reuse was selected as the dependent variable, and the system-, information-, interface-, and service-quality were selected as independent variables based on a previous study. Additionally, the fields, grades, and experience were selected as moderator variables, and an analysis was conducted. The result of this study can be used to establish an EDISON platform development strategy that considers user support in terms of fields, experience, and grades.

***Keywords*:** Computational Science and Engineering, EDISON platform, Moderating Effect, Survey

1. Introduction

The EDISON platform is the only computational science and engineering platform in Korea that was developed and has been deployed since 2011. The platform has been providing services to around 900 types of software and 850 types of contents in seven fields: computational fluid dynamics, nanophysics, computational chemistry, computational structural dynamics, computer-aided optimal design, computational medicine, and computational urban environment [1,2]. As of October 2020, 2,286 courses in 59 universities has been accessed by 79,000 cumulative users. Most users of the EDISON platform are science and engineering university students who use the platform for educational and research purposes.

This research was supported by the KISTI Program (No. K-20-L02-C05-S01), the EDISON Program through the National Research Foundation of Korea (NRF) (No. NRF-2011-0020576), and the NRF and the Center for Women In Science, Engineering and Technology (WISET-2020-190).

The users on the EDISON platform partakes in surveys that have been conducted for many years, and several analysis studies have been conducted on the factors that influence the reuse of the platform. However, no research has yet been conducted on how the background of users affects their intention to reuse (IR) the platform. In this study, we analyze how the experience in computational science engineering, grades, and fields of study affect the user's intention to reuse the EDISON platform. Further, this study will contribute to the establishment of an operational platform strategy that considers the background of the platform users.

2. Research Model and Research Method

In previous studies, it was observed that the quality factor of the computational science and engineering platform affects the user's IR [3]. Accordingly, in this study, IR was used as a dependent variable. The four factors that affect IR: system quality (SYSQ), information quality (INFQ), interface quality (INTQ), and service quality (SERQ), were used as independent variables. In this study, fields (Field), grades (Grade), and experience in computational science engineering (Exp) were used as moderator variables to determine how the background of users affects the user's IR in terms of the platform's quality factor.

3. Empirical-Analysis Results

This study analyzed the effect of fields, grades, and experience on the user's intention to reuse the platform. The users were participants in the EDISON SW Utilization Contest, which intensively utilizes the EDISON platform. To perform an analysis, a survey was conducted on 510 participants of the EDISON SW Utilization Contest over three years from 2018 to 2020. **Table 1** lists the demographic analysis of the 510 participants in the EDISON SW Utilization Contest.

Table 1. Results of the demographic analysis

Item	Classification	Frequency	Percentage (%)
Year	2018	182	35.7
	2019	175	34.3
	2020	153	30.0
Grade	First year	14	2.7
	Second year	21	4.1
	Third year	105	20.6
	Fourth year	140	27.5
	Master course	150	29.4
	Doctoral course	80	15.7
Field	Computational fluid dynamics	33	6.5
	Nanophysics	21	4.1
	Computational chemistry	105	20.6
	Computational structural dynamics	140	27.5
	Computer-aided optimal design	150	29.4
	Computational medicine	80	15.7
	Computational urban environment	78	15.3
Experience	None	107	21.0
	<6 months	175	34.3
	6 months–1 year	82	16.1
	1–3 years	100	19.6
	3–5 years	26	5.1
	>5 years	20	3.9

A confirmatory factor analysis was performed to verify the validity of the questionnaire used in the survey. **Table 2** lists that the p value of all questionnaires was 0.001 or less, and the standardized factor loading value was 0.6 or more. According to this factor loading value, all factors were more than 0.7 based on the results obtained by calculating the construct reliability (CR) of the measurement items for each variable. Additionally, we observe that all factors were more than 0.5 when the average variance extracted (AVE) value was calculated. Therefore, the convergent validity for evaluating the consistency of the variables was confirmed in reference [4].

Table 2. Results of the confirmatory factor analysis

Factor	Indicator	Estimate		SE	p	Stand. Estimate	CR	AVE
SYSQ	sysq1	1.37		0.0581	<0.001	0.863	0.879	0.596
	sysq2	1.35		0.0551	<0.001	0.881		
	sysq3	1.10		0.0597	<0.001	0.730		
	sysq4	1.04		0.0644	<0.001	0.665		
	sysq5	1.12		0.0649	<0.001	0.696		
INFQ	infq1	1.15		0.0514	<0.001	0.820	0.942	0.698
	infq2	1.18		0.0530	<0.001	0.816		
	infq3	1.16		0.0494	<0.001	0.847		
	infq4	1.17		0.0495	<0.001	0.851		
	infq5	1.21		0.0497	<0.001	0.864		
	infq6	1.13		0.0476	<0.001	0.853		
	infq7	1.09		0.0508	<0.001	0.796		
INTQ	intq1	1.26		0.0569	<0.001	0.815	0.931	0.729
	intq2	1.37		0.0572	<0.001	0.857		
	intq3	1.36		0.0539	<0.001	0.888		
	intq4	1.35		0.0525	<0.001	0.896		
	intq5	1.25		0.0575	<0.001	0.808		
SERQ	serq1	1.29		0.0511	<0.001	0.883	0.956	0.814
	serq2	1.20		0.0457	<0.001	0.902		
	serq3	1.30		0.0454	<0.001	0.948		
	serq4	1.26		0.0458	<0.001	0.925		
	serq5	1.23		0.0520	<0.001	0.849		
IR	ir1	1.39		0.0519	<0.001	0.915	0.951	0.829
	ir2	1.43		0.0529	<0.001	0.917		
	ir3	1.36		0.0536	<0.001	0.884		
	ir4	1.34		0.0490	<0.001	0.925		

Table 3 indicates that the grades did not have a moderating effect between all quality factors and IR.

Table 3. Analysis results of the moderating effect of the grades

Factor	Classification	Estimate	SE	p
SYSQ	IR	0.6830	0.0310	<0.001
	Grade	0.0387	0.0246	0.115
	IR * Grade	0.0142	0.0242	0.557
INFQ	IR	0.8101	0.0265	<0.001
	Grade	−0.0245	0.0210	0.243
	IR * Grade	0.0389	0.0207	0.060
INTQ	IR	0.6415	0.0341	<0.001
	Grade	−0.0256	0.0271	0.345
	IR * Grade	0.0248	0.0266	0.353
SERQ	IR	0.71628	0.0315	<0.001
	Grade	0.00589	0.0250	0.814
	IR * Grade	0.01405	0.0246	0.568

In **Table 4**, the field demonstrated a pure moderator effect wherein the field interacted with the INFQ which did not show any direct influence on the IR. Furthermore, the field had a quasi-moderating effect wherein it interacted with the SERQ and induced a direct influence on the IR [5,6].

Table 4. Results of the analysis of the moderating effect of the fields

Factor	Classification	Estimate	SE	p
SYSQ	IR	0.6917	0.0306	<0.001
	Field	−0.0344	0.0167	0.040
	IR * Field	0.0338	0.0186	0.069
INFQ	IR	0.7987	0.0262	<0.001
	Field	−0.0154	0.0143	0.284
	IR * Field	0.0336	0.0159	0.035
INTQ	IR	0.6319	0.0338	<0.001
	Field	−0.0248	0.0185	0.179
	IR * Field	0.0214	0.0205	0.296
SERQ	IR	0.7175	0.0308	<0.001
	Field	−0.0545	0.0169	0.001
	IR * Field	0.0464	0.0187	0.013

In **Table 5**, we observe that the experience in computational science engineering did not have any moderating effect between all the quality factors and IR.

Table 5. Results of the moderating-effect analysis according to experience

Factor	Classification	Estimate	SE	p
SYSQ	IR	0.67971	0.0306	<0.001
	Exp	0.06184	0.0221	0.005
	IR * Exp	−0.00623	0.0231	0.788
INFQ	IR	0.7916	0.0263	<0.001
	Exp	0.0287	0.0190	0.132
	IR * Exp	0.0127	0.0199	0.522
INTQ	IR	0.63388	0.0338	<0.001
	Exp	−0.00120	0.0244	0.961
	IR * Exp	−0.03907	0.0256	0.126
SERQ	IR	0.7121	0.0312	<0.001
	Exp	0.0289	0.0226	0.201
	IR * Exp	-0.0261	0.0236	0.269

4. Summary and Future Works

This study aims to determine how the background of users in computational science and engineering platform affects their intention to reuse the platform. Hitherto, we analyzed the moderating effects of the field, grade, and experience on the platform between the quality factors of the platform and IR. To perform the analysis, a survey was conducted on the participants in the EDISON SW Utilization Contest.

The analysis suggested that the field demonstrated a significant influence on the IR by interacting with the INFQ and exerted a significant influence on the IR by interacting with the SERQ. On the other hand, the grade and experience in computational science engineering did not show any moderating effect.

The results obtained suggest that collecting and responding to the user's requirements in each field is essential to improve the information and service qualities of the platform. The results obtained from this study will contribute to the establishment of an EDISON platform development strategy in the future.

References

[1] Y. K. Suh, H. Ryu, H. G. Kim, and K. W. Cho, "Edison: A Web-Based HPC Simulation Execution Framework for Large-Scale Scientific Computing Software,", 2016 *16th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, 608-612, May 2016.

[2] J. Ma, J. R. Lee, K. W. Cho, and M. J. Park, "Design and Implementation of Information Management Tools for the Edison Open Platform," *KSII Transactions on Internet and Information Systems*, Vol. 11, No. 2, pp. 1089-1110, 2017.

[3] N. On, G. Ryu, M. Koh, J. R. Lee, and N. Kim, "An Empirical Study on the Intention to Reuse Computational Science and Engineering Platforms: A Case Study of Edison," *KSII Transactions on Internet and Information Systems*, Vol. 14, No. 8, pp. 3437-3456, 2020.

[4] W. W. Chin, "The Partial Least Squares Approach for Structural Equation Modeling. In Methodology for business and management *Modern methods for business research*, p G. A. Marcoulides (Ed.), Lawrence Erlbaum Associates Publishers, 1998, 295-336.

[5] A. L. Darrow, and D. R. Kahl, "A Comparison of Moderated Regression Techniques Considering Strength of Effect," *Journal of Management*, Vol. 8, No. 2, pp. 35-47, 1982.

[6] S. Sharma, R. M. Durand, and O. Gur-Arie, "Identification and Analysis of Moderator Variables," *Journal of Marketing Research*, Vol. 18, No. 3, pp. 291-300, 1981.

A Study on Satisfaction of Users in Computational Science and Engineering Platform - The Case of University

Noori On, INHO JEON, Namgyu Kim, and Jongsuk Ruth Lee*
Division of National Supercomputing, Korea Institute of Science and Technology Information
Daejeon, South Korea
[Email: claireon, inojeon, ssgyu, jsruthlee@kisti.re.kr]
*Corresponding author: Jongsuk Ruth Lee

Abstract

The Korea Institute of Science and Technology Information (KISTI) created a computational science and engineering platform (the EDISON Platform) through the EDISON project in order to expand the base of computational science engineering and develop simulation software (SW) and technology for computational science engineering. There is little research on how technologies in the EDISON platform, developed over nearly ten years, affect the satisfaction of real users.

In this study, students using EDISON simulation SW in university classes are surveyed, and the effect of EDISON simulation software utilization education on user satisfaction is analyzed. The survey was conducted with students who took classes using the EDISON platform in the first semester of 2020. The survey of actual learners showed that the use of simulation platforms in class produced a positive response. The results of this study lead to strategic implications for establishing improvement plans for the computational science and engineering platform.

***Keywords*:** computational science and engineering, computational science and engineering platform, EDISON, simulation software

1. Introduction

The field of computational science, which combines basic research and computing technology, is bringing about positive results by innovating scholarly research. The Korea Institute of Science and Technology Information has been working on the EDISON project since 2011 to expand the base of computational science in Korea and to commercialize computational science simulation software. The EDISON platform operating in the business is equipped with simulation software to enable research in seven fields of science and engineering on the web. In this study, a survey is conducted with students using EDISON simulation software in university classes, and the effect of EDISON simulation software utilization education on learners' satisfaction is analyzed. The survey was conducted with students who took courses using the EDISON platform during the first semester of 2020. This study evaluated the status of the computational science platform from the user's perspective and based on user satisfaction.

2. The EDISON Project

Basic science fields such as physics, chemistry, and material science combine supercomputers

This research was supported by the KISTI Program (No. K-20-L02-C05-S01), the EDISON Program through the National Research Foundation of Korea (NRF) (No. NRF-2011-0020576), and the NRF and the Center for Women In Science, Engineering and Technology (WISET-2020-190).

and high-performance network-based cyber infrastructure with the sciences of calculation and engineering. In recent years, the field of computational science and engineering has increased and expanded HPC utilization due to rapid advances in computing performance, networking, and storage [1]. Computational science and engineering are being used in various fields for educational and research purposes [2]. The correct image segmentation provides clinically important information to analyze the tissue for physicians as well as to reconstruct 3D volume model based on the segmentation result.

Since 2011, the Korea Institution of Science and Technology Information (KISTI) has been carrying out the EDucation-research Integration through Simulation On the Net (EDISON) project to expand the base of Korea's computational science engineering and to foster human resources in Korea's computational science engineering field. The EDISON Project consists of a central center as well as specialized centers in seven fields: computational fluid dynamics, computational nanophysics, computational chemistry, computational structural dynamics, computer aided optimal design, computational medicine, and urban environment. Specialized centers develop and mount simulation software programs in the science and engineering fields, utilizing them in undergraduate and graduate classes. The central center develops and serves the platform's core technologies that can be equipped with simulation software made at specialized centers. It also supports the community in order to expand the base of computing science. The EDISON platform currently utilizes 900 types of software, 848 cases of content, and making use of approximately 75,000 people for 9 years

3. The Empirical Analysis

This study surveyed students who took the EDISON platform utilization class during the first semester of 2020. Analysis was performed based on 1,379 responses. Empirical analysis was conducted through the Jamovi program [3, 4, 5].

According to survey queries about the respondents' fields of study, 85% of respondents were in the field of computational chemistry, with others studying computational fluid dynamics (6.9%), computational medicine (3.0%), urban environment (0.7%), and computational nanophysics (0.1%).

Most respondents (91.4%) had no experience using simulation software. According to survey queries about the most important factors in software utilization classes, 48.9% of respondents named experience and guidance as the most important factors in simulation software utilization classes; 37.6% of respondents reported that practical programs and related content that provided ease of use for the software experience were most important.

Regression analysis was conducted on the satisfaction of users using the EDISON platform and incorporated five factors. An easy-to-understand and clear web portal User Interface(UI), an easy-to-learn web portal usage method, an appropriate practice environment, a sufficient explanation from the lecturer that enabled smooth software use, and an easy-to-understand software instruction manual as well as lecture materials were selected and analyzed as factors of platform utilization satisfaction.

The F value of the regression model is 376, with a significance probability of <0.001, which is statistically significant at .001, and the independent variables account for 57.8% of the total variance of user satisfaction. ANOVA analysis was performed to check whether each factor used as the independent variable in the regression model had satisfaction and statistical significance on the dependent variable. It was concluded that all factors could be interpreted as factors that statistically and significantly explain satisfaction at the significance level of .05. In addition, it was confirmed that the VIF values of the independent variables were all 10 or less; therefore, there was no multicollinearity [6].

Among all of the factors that had an effect on satisfaction at the significance level of .05, it was confirmed that the easy-to-understand and clear web portal UI (0.1956) had the most impact. Appropriate practice environment (0.1943), easy-to-learn web portal usage (0.1898), easy-to-understand software instruction manual and lecture data (0.1384), and sufficient explanation by lecturer (0.0893) also had significance.

4. Conclusions

In this study, an analysis was conducted on user satisfaction with the EDISON computational science platform. Survey respondents were university students participating in classes using the EDISON platform during the first semester of 2020. The study confirmed that the user interface of the easy-to-understand web portal and the user manual and lecture data of the easy-to-understand software affected users' satisfaction. This result indicates that providing an environment that users can employ with ease was more important than the technical performance of the platform.

The limitation of this study is that most survey respondents had no previous simulation software utilization experience. While this study revealed the importance of providing a smooth foundation and environment for novice users, it remains necessary to identify factors that users have used the platform consistently consider essential to platform satisfaction.

Based on this study, future study will be conducted with computational scientists and those results compared with this study. The studies are likely to contribute to the establishment of technology development strategy for the computational science and engineering platform.

References

[1] Y. K. Suh, H. Ryu, H. G. Kim, and K. W. Cho, "Edison: A Web-Based HPC Simulation Execution Framework for Large-Scale Scientific Computing Software," (2016) *16th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, Cartagena*, pp. 608-612, 2016.

[2] J. H. Gwak, "An Introduction to a Computational Science Cooperative Course," *Natural Science*, 2004 fall *issue*, vol. 17, page 227.

[3] J. H. Gwak, "An Introduction to a Computational Science Cooperative Course," *Natural Science*, 2004 fall *issue*, vol. 17, page 227.

[4] R Core Team (2019). R: A Language and environment for statistical computing. (Version 3.6) [Computer software]. Retrieved from https://cran.r-project.org/.

[5] R Core Team (2019). R: A Language and environment for statistical computing. (Version 3.6) [Computer software]. Retrieved from https://cran.r-project.org/.

[6] T. S. Kim, J. H. Jang, and P. G. Back, "Sanriojungsimjamovitonggyebunsuk," *Parkyoungstory*, 2020

Automation technique of parallel distributed processing for machine learning on HPC-based AI platform

Seokkeun Yi
National Institute of Supercomputing and Networking Center for Computational Science Platform,
Korea Institute of Science and Technology Information
Daejeon, South Korea
[e-mail: sky@kisti.re.kr]
*Corresponding author: Seokkeun Yi

Abstract

With the active flow of AI research, large IT companies have launched many AI services. Many of these services offer a wide range of equipment choices, such as high-performance GPUs and offer a wide range of AI services based on the cloud. However, when AI users deal with big data, it is difficult to write parallel/distributed codes directly based on understanding the environment of the cloud in order to handle parallel/distributed processing in these services. On the contrary, there are intuitive GUI-based AI cloud services that simplify this process. However there is a limit that it is difficult to modify or implement internal code in this case. Thus, in this study, we researched a technique that allows AI developers or data engineers who are difficult to perform parallel/distributed operations to easily perform parallel/distributed processing with simple coding. With this technique, a Jupyter notebook can be automatically distributed parallelized and executed with simple API function call by creating neural network or machine learning algorithm and pre-processing data in the notebook environment. It also provides an API to monitor training job. The technique was implemented as an IDE called 'AI Studio' in 'EDISON AI Platform', a convenient integrated development environment. This IDE makes it easy to perform AI workflow such as uploading/managing data, running and inspecting training jobs, and managing and deploying models.

***Keywords*:** AI, Deep Learning, Machine Learning, AI Platform, ML Platform, Artificial Intelligence, Parallel Distributed Processing, Parallel Distributed Processing Automation, IDE

1. Introduction

Thanks to intense interest in AI, development of AI technology is gradually accelerating. According to this flow, large IT companies are offering various AI services. Among them, AI platforms have been developed so that AI researchers can utilize them. Representative services are Amazon's Sagemaker [1], Microsoft's Azure Machine Learning [2], Google's AI Platform and IBM's Watson [3]. These platforms support various deep learning/machine learning frameworks (e.g. Tensorflow, Keras, PyTorch, Scikit-Learn, etc.) for research in various fields and provide a high level cloud service. It also provides high-performance GPU resources for fast computing. These AI platforms have a wide range of features and resources, but there are also challenges. It's complexity. Users have to search through a lot of documents to do what they want, and the many features make it difficult to use without education. It is also very difficult for

users to write their own code because they need to have a high understanding of parallel/distributed processing in order to write a machine learning code that performs parallel/ distributed operations. On the contrary, there are intuitive GUI-based AI cloud services that simplify this process. However there is a limit that it is difficult to modify or implement internal code in this case. Therefore, in this study, an integrated development environment (IDE) was built on the AI platform so that researchers who understand artificial intelligence to some extent can flexibly modify the code and perform artificial intelligence research and training easily in a GUI environment. In particular, in this process, parallel/distributed automation technique was introduced to solve the difficulty of writing parallel/distributed code, which many artificial intelligence researchers are facing. This IDE allows users to easily perform parallel/ distributed machine learning training jobs using freely written code. In addition, it has been developed so that all these series of tasks can be performed on a HPC machine through a simple application programming interface (API) function call without any knowledge of complicated HPC task management.

2. Integrated Development Environment on EDISON AI Platform

2.1 EDISON AI Platform

KISTI operates a computational science platform called EDISON (EDucation-research Integration through Simulation On the Platform and Net), which is the bridge to computational science for education and advanced research. EDISON is Korea's only integrated R&ED platform that provides web-based services such as simulation SWs and AI analysis in various fields using KISTI's computing resources. Among various fields, a platform for AI is also being serviced, and in this platform called 'EDISON AI Platform'. However, since the service is specialized in a specific field, it is difficult to be used universally. So, a general-purpose development environment has been established called ‘AI Studio’ on EDISON AI Platform.

2.2 AI Studio (IDE)

As an integrated development environment (IDE), AI Studio has been developed for AI researchers to easily develop machine learning codes and research AI. AI Studio allows users to create their own workspaces, and the dashboard makes it easy to manage data engineering jobs or AI training jobs running in the workspace. It is a powerful IDE that allows easy dataset processing/management, Jupyter notebook work, training job monitoring/management, model management, and service deployment following common AI workflow. In addition, a user-friendly web service is provided. In particular, this study provides an automation technique of parallel/distributed processing that helps users to write parallel/ distributed code automatically with only basic AI knowledge in notebook environment. In addition, we develop and provide an API that can easily perform machine learning training jobs and data preprocessing on HPC. The automatically parallel/distributed code is easily allocated to resources by using API functions without any knowledge of HPC. With this simple usage, parallel/distributed jobs are allocated resources and delivered to queues, which are sequentially processed through HPC's job scheduler, KISTI's computational resource.

Fig. 1. AI Studio in EDISON AI Platform

3. Automation technique of parallel distributed processing for machine learning

3.1 Machine learning training workflow

At first, to do parallel/distributed training job of machine learning in AI Studio, users need to run

a notebook instance through the studio. When the notebook is running, users need to load the estimator module according to the framework used for deep learning or machine learning. With the estimator module loaded, it is now needed to write an algorithm (e.g. linear regression) or neural network for training of machine learning. And next, writing code for data preprocessing has to be done. Here, users can freely write code for data preprocessing such as normalization or dropout.

The class used to make the neural network or data preprocessing is provided as a template, and only this part of the code need to be written in that notebook template. When the algorithm/neural network and data preprocessing code are completed, various parameters required for training can be entered. Once it has been done so far, parallel/distributed training is all ready to do. Now, just call the API function used for machine learning training, and the parallel/distributed machine learning code which uses Horovod [4] is automatically generated based on the contents of the notebook and training is performed. At this time, an API function for monitoring the training job is also provided, so if the function is called, the graph that the training is progressing in real time is shown. This training job is performed on KISTI's HPC resource (a resource separate from the web service) connected to the AI Studio service, and the job sent to this HPC is queued through the job sheduler according to the HPC infrastructure. Users do not need to write any code related with HPC at all. Similarly, when canceling the requested training job, training job can be canceled by simply calling the API function for cancel. When the training job is complete and a user like the performance of the model, the user can register the model so that it can be deployed as a service from a web service. Similarly, this can be done by simply calling a simple API function. In this way, AI Studio of the EDISON AI Platform has implemented automation technique of parallel/distributed processing. In other words, after simply writing the algorithm and data preprocessing code in the notebook environment, you can easily proceed parallel / distributed machine learning training in HPC by simply submitting the API function provided by this AI Studio.

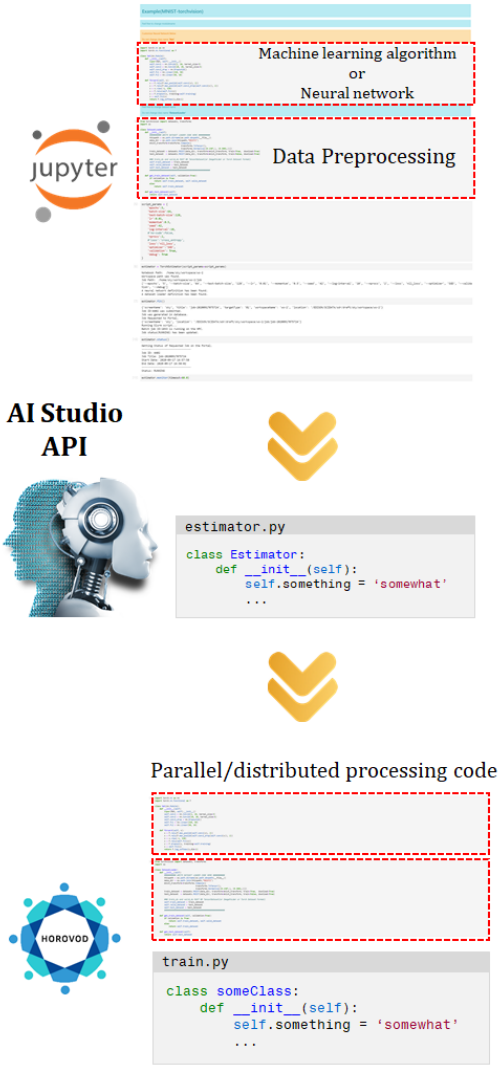


Fig. 2. Machine learning training workflow in AI Studio

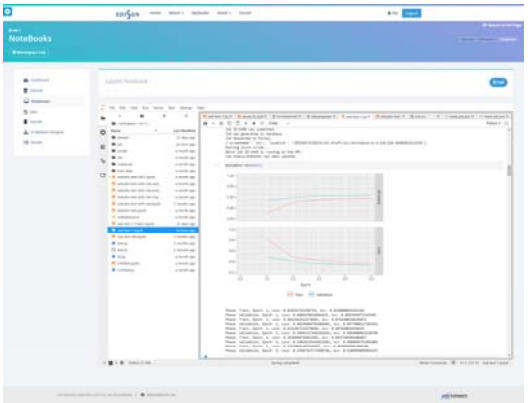


Fig. 3. Training job monitoring in the notebook

3.2 Inference with trained model

It is also very simple to make inferences using the trained model. As in Section 3.1, first create a new notebook in the studio. And load the estimator module of the desired machine learning framework. At this time, it is only needed to enter the name of the model to be used (registered earlier). After that, the only code that the user will write is the data preprocessing code necessary for inference. When this code is written, only some parameters for parallel/ distributed operation and hyperparamters for training are needed. When the inference API function is called, and the prediction job is requested to HPC through the automatically generated parallel/distributed code. And the status and results of inference job can be checked through simple API functions. Of course, canceling a job can be done in the same way as training.

4. Discussions

This study focused on the automation technique of parallel distributed processing for machine learning. In addition, this automation technique will be effective in lowering the barriers to entry for AI researchers to the large data processing technology required for big data research. Nevertheless, there are some limitations in solving problems in various cases. Only basic types of hyper-parameters are inputted as training parameters, and loss functions or optimizers are dependent on the provided framework functions (e.g. PyTorch). Also, there is a structural limitation of the automatically generated machine learning code. Since a single algorithm class is used, the difficulty of code development may increase even further to request complex training tasks. And another limitation of this technology is that it is somewhat difficult for a complete beginner to use it because it is necessary to write a machine learning algorithm or preprocessing code with basic AI knowledge based on understanding of basic python syntax. In the future, it is planned to increase the type of input training parameter input and the support of the machine learning framework and improve the technology to solve complex artificial intelligence problems.

5. Conclusions

Despite the limitations discussed above, it is more flexible, easy to use, and effective for handling big data compared to other commercial AI platforms. And it is more suitable for beginners and intermediates who have difficulty in parallel/distributed coding. In other words, it is a great advantage of this technology that if a user simply write a machine learning algorithm or neural network and write data preprocessing code, it is easy to perform parallel/distributed work without any knowledge of HPC. It will also be a good tool to increase the productivity of AI researchers dealing with big data in that it fully automates parallel/distributed tasks.

Acknowledgement

This research was supported by the KISTI Program (No. K-20-L02-C05-S01), the EDISON Program through the National Research Foundation of Korea (NRF) (No. NRF-2011-0020576). A Grant was awarded by the Ministry of Science and ICT (MSIT) under the Program for Returners for R&D.

References

[1] E. Liberty, Z. Karnin, B. Xiang, L. Rouesnel, B. Coskun, R. Nallapati, ... and C. Balioglu, “Elastic Machine Learning Algorithms in Amazon SageMaker,” in *Proc. of the 2020 ACM SIGMOD International Conference on Management of Data*, pp. 731-737, 2020.

[2] J. Barnes, “Azure machine learning,” *Microsoft Azure Essentials. 1st ed(Microsoft)* , 2015.

[3] Y. Chen, J. E. Argentinis, and G. Weber, “IBM Watson: how cognitive computing can be applied to big data challenges in life sciences research,” *Clinical therapeutics*, Vol 38, pp. 688-701, 2016.

[4] A. Sergeev, and M. Del Balso, “Horovod: fast and easy distributed deep learning in TensorFlow,” *arXiv preprint arXiv*, 1802.05799, 2018.

Data science education improvement plan based on demand survey

Haejin Kang and Seunghwa Jung
[Data science school, Korea Institute of Science and Technology Information,
Seoul, South Korea
Depr. Of Liveral Arts & Interdisciplinary Studies, Hanseo University,
Seosan, South Korea
[e-mail: edu_kang@kisti.re.kr, edutech@hanseo.ac.kr]
*Corresponding author: Seunghwa Jung

Abstract

This study aims to analyze science and technology experts' level of interest in data science education and examine their demand for such education. A survey of members of the National Digital Science Library (NDSL) website was conducted for a month**,** between August and September 2018, and 1,836 members responded to it. Survey questions based on 48 subjects in ten fields related to data science were included to evaluate respondents' interest fields, and multiple choices were allowed for these questions. The survey results showed that the most popular subject was Basic information retrieval (71.90%), followed by survey, analysis, and R&D research and analysis & selection strategy (67.6%), and Excel (64.20%). The results of examining education course types indicated that respondents preferred practical training (85.40%) to theoretical courses (69.17%). When both practical training and a theoretical course were combined, most respondents were willing to take them combined based on the ratio of 1:1. As for scale of education, a small-sized course (59.48%), including 15 students or less, was the most preferred. Moreover, 59.23% of the respondents preferred offline courses, whereas 40.72% of them were more willing to take online courses. The results of analyzing respondents' demand for data science education implies that it is necessary to increase data application education courses and develop short-term online education courses for enhancing learner capabilities.

***Keywords*:** Data science education, HPC, Data education,

1. Introduction

The Korea Institute of Science and Technology Information (KISTI) is a government-sponsored research organization responsible for providing professional education on data use and high performance computing (HPC) for those who work in industrial, academic, institutional, and organizational sectors through the application of its infrastructure and labor force. To this end, the KISTI has been operating education courses on data use and HPC since 2001. However, the need for adjusting existing education courses and content has been raised due to the transition of Korean society into the fourth industrial revolution era.

Adult learners account for most of the workers who are selected as subjects for the KISTI education courses and belong to certain types of occupations. They tend to participate in education courses only when they can achieve goals given an adequate amount of time and cost. For this reason, the necessity of providing educational content and methods reflecting

This work was supported by KISTI Program(K-20-L05-C01-S01) and the National Research Foundation of Korea(NRF) and the Center for Women In Science, Engineering and Technology(WISET) Grant funded by the Ministry of Science and ICT(MSIT) under the Program for Returners into R&D in 2020 (WISET-2020-225).

learners' learning needs has been consistently highlighted [1-3].

The KISTI performed a survey on learner demand for developing their favorite education courses. The survey target was limited to NDSL members, a service platform for providing scientific and technological information to operate education courses, in consideration of the professional labor force and infrastructure secured by the KISTI. The NDSL website is frequently used by science and technology experts, given that it provides a database search for over 100 million documents, mainly based on science and technology, according to specific options such as Korean theses, patents, reports, trends, journals, proceedings, researchers, and research institutes.

The survey was carried out to examine science and technology experts' level of interest in existing education courses and new subjects to be added. Their intention in attending professional data science education courses, reasons for attending such courses, and preferred types of education were also examined to identify their educational content demand and review the existing educational operation methods.

2. Research Method

The Web-based survey of demand was made available to the approximately 80,000 members of the NDSL website, from Aug 29th to Sept 20th, 2018. Structured questionnaires were provided to the survey subjects, and 1,836 subjects responded to the questionnaires. The response rate based on the number of valid respondents was calculated to be approximately 2.3%. Fundamental questions for respondents included basic personal information, level of interest for data science education courses, intention of attending data science education courses provided by the KISTI, need for the advertisement of such courses, preferred types and methods of education, appropriate cost for education, and activities of instructors. In this study, only a few necessary questions among them were extracted and used for analysis.

Respondents were allowed multiple choices for their favorite education courses among 48 education courses in ten fields to investigate their level of interest in data science education courses. The KISTI was already providing 21 courses (12 courses on data use and nine courses on HPC) among the 48 courses indicated as examples in the questionnaire at the time of the survey. The other 27 courses were based on the subjects that the KISTI had proposed to prepare for the fourth industrial revolution era. **Table 1** shows fields and education courses indicated as examples in the questionnaire.

Table 1. Information of respondent

Category	Division	No.of participants	Ratio (%)
Job	undergraduate student	125	6.8
	graduate student	162	8.8
	professor	183	10.0
	researcher	613	33.4
	employee	604	32.9
	etc.	149	8.1
Age	20s	310	16.9
	30s	571	31.1
	40s	586	31.9
	50s	312	17.0
	60s or older	57	3.1
Career year for data science	less than 5 years	384	20.9
	more than 5 years	484	26.4
	no experience	763	41.6

Table 2. Information of respondent

field	Curriculum
Information retrieval	Basic information retrieval* license information search*
Information Analysis	Industrial market research and analysis * Research and analysis* Technology roadmap * Patent map expert training* Future technology prediction* Data Management Project
R&D	R&D plan · management · assessment* R&D research and analysis & selection strategy*
Technology transfer	Technology transfer commercialization* Practical market research for planning Technology commercialization* Calculation of technical fees & negotiation tactics*
Programming	R Python** C language JAVA Fortran** Web programing Platform (SPRING, LIFELAY etc.)
Basic theory	Basic mathematics Numerical analysis ** Deep learning Voice recognition AI

	Data base theory Statistics theory
Data processing	Hadoop Spark NoSQL Excel
HPC	5th supercomputer** MPI** OpenMP** Hybrid(openMP+MPI) ** Cloud Computational science Application software (ABAQUS, FLUENT etc.) Optimization** GPU**
IoT	Android Arduino Raspberry Pi Drone
etc.	Protection of Information & security Network Virtual reality Mobile application

*existing DATA Curriculum

**existing HPC Curriculum

A frequency analysis of survey results was conducted based on Excel, and the percent of cases were presented for questions with multiple choices allowed.

3. Results

3.1 Intention of attending professional data science education courses provided by the KISTI

A total of 82.19% of all respondents were willing to attend professional data science education courses provided by the KISTI, accounting for a high proportion. Regarding reasons for attending data science education courses, those related to improvement of personal capabilities, such as an increase in job performance (65.03%) and fundamental capacity development (52.91%), were selected the most. It was also found that many respondents were willing to utilize the advantages of KISTI, such as the use of KISTI's infrastructure (30.66%) and interactions with experts related to the KISTI (29.14%). In addition, 17.81% of all respondents responded that they would not attend professional data science education courses provided by the KISTI. Regarding reasons for not attending these courses, insufficient time (62.20%) was selected the most, followed by an unrelated major or job (36.59%) and the burden of a tuition fee (28.05%).

3.2 Preferred types of education

Respondents were allowed to make multiple choices for preferred types of education, and it was found that they were highly eager to take practical training (85.40%) and theoretical lectures (69.17%). When theoretical lectures and practical training were combined, the ratio of 5:5 for operating such combined course was regarded as the most appropriate ratio, followed by those of 6:4 (22.77%) and 7:3 (18.68%). Meaning, more respondents demanded the combination of theoretical lectures and practical training based on the ratio of 5:5 or those with a slightly greater proportion of theoretical lectures. With regard to the preferred number of students in a course, a small-sized course including 15 students or less (59.48%) was selected the most, followed by a middle-sized course including 16 to 35 students (37.47%), and a large-sized course including 36 students or more (3.05%). This result indicated that most respondents preferred a small-sized course.

Regarding questions on the preferred type of education courses between online and offline, 59.23% of all respondents preferred offline courses, whereas 40.72% preferred online courses. Moreover, a preferred time slot and a preferred amount of time for attending education courses between respondents who preferred online courses and those who preferred offline courses were compared. The comparison result indicated that those who preferred online courses felt more comfortable attending education courses during off-hours, such as the weekend and in the evening on weekdays, than those who preferred offline courses. The former also tended to prefer shorter duration courses to courses requiring four hours or more.

Table 3. a comparison of preferred time slot

	on-line(N=748)		off-line(N=1088)	
	No. of participants	**case percent**	**No. of participants**	**case percent**
weekday morning	135	18.05%	349	32.14%
weekday afternoon	223	29.81%	484	44.57%
weeknight	299	39.97%	346	31.86%
weekend morning	258	34.49%	350	32.23%
weekend afternoon	213	28.48%	259	23.85%
weekend evening	107	14.30%	133	12.25%

Table 4. a comparison of amount of time for attending education courses

	on-line(N=748)		off-line(N=1088)	
	No. of participants	**case percent**	**No. of participants**	**case percent**
~ 2 hours	218	29.1%	157	14.42%
2 ~ 4 hours	329	44.0%	472	43.34%
4 ~ 6 hours	140	18.7%	283	25.99%
6 ~ 8 hours	49	6.6%	154	14.14%
8 hours ~	12	1.6%	22	2.02%

3.3 Level of interest in education courses

Respondents were allowed to make multiple choices for their favorite education courses among 48 courses in ten fields. The KISTI was already providing 21 courses (12 courses on data use and nine courses on HPC) among the 48 courses indicated as examples in the questionnaire at the time of the survey. The other 27 courses were included as they were estimated to be demanded by potential learners.

It was found that basic information retrieval received the most attention from respondents among the educational courses, and 71.9% of all respondents were interested in this course. The following table presents the list of education courses in which more than 50% of all respondents were interested. As for subjects not included in existing educational courses, respondents showed a high level of interest in Excel, Android, artificial intelligence (AI), and statistical theories. Regarding existing HPC education courses, 32.63% of all respondents were interested in numerical analysis and 28.05% in Python. However, the level of interest for other subjects besides them was calculated to be below 15%. Among the HPC-related subjects estimated to be demanded by potential learners, those which showed a level of interest of 20% or higher included cloud computing (45.32%), applied software such as ABAQUS and FLUENT (23.69%), and computational science (20.32%).

Table 5. the list of education courses in which more than 50% of all respondents were interested

Division	**Curriculum**	**case percent**
data	Basic information retrieval	71.90%
data	R&D research and analysis & selection strategy	67.59%
new	Excel	64.16%
data	Practical market research for planning Technology commercialization	56.81%
new	Android	55.61%
data	Future technology prediction	54.52%
data	license information search	53.81%
data	R&D plan · management · assessment	53.38%
new	AI	51.58%
data	Industrial market research and analysis	51.31%
new	Statistics theory	50.98%

DATA = existing DATA Curriculum
HPC = existing HPC Curriculum
NEW = topic for new Curriculum

4. Conclusion

4.1 Enhancement of existing educational content based on data science with the advent of the bigdata era

The skills required for major talent who will lead the 21st century were established as the core skills of the 21st century learners to prepare for the future society. Moreover, information access, evaluation, and use were regarded as necessary abilities to increase information literacy[4]. In modern society, people can easily search an enormous amount of knowledge and perform business transactions, networking, and learning online. However, as the online environment has both advantages and disadvantages, people should increase their information literacy capabilities to use online information more effectively.

The education course of basic information search has been operated by the KISTI for 18 years and was canceled in 2019 due to a decrease in the number of applicants. The high level of interest in this course found in this study implies that potential learners are willing to increase their information literacy capabilities, as indicated above. When this course was launched for the first time in 2001, basic information search simply referred to a method of searching information. However, this course's content should be restructured according to the core skills of 21st century learners.

Education courses which showed a level of interest of 50% or higher mainly consisted of those on using information to derive results (i.e. survey, analysis, and selection strategies for R&D subjects, practical market investigation for

technology project planning, future technology prediction, and industrial market investigation and analysis) and those on practically using data (i.e. Excel, AI, and statistical theories). Adult learners tend to perform self-directed learning for education courses that they selected themselves and participate in such learning activities to achieve practical purposes, such as skill acquisition, rather than personal development [5,6]. Thus, the KISTI should develop practical education courses which allow learners to use and apply data by themselves, beyond existing simple information search content, to satisfy the needs of adult learners who are the main targets for the KISTI education courses.

4.2 Necessity of online education course development

According to the survey analysis results, more than 40% of all respondents preferred online education courses, although the number of those who preferred offline education courses was higher than that of those who preferred online education courses. It was also found that those who preferred online education were willing to attend short-duration education courses, requiring less than four hours, during off-hours. Therefore, it is anticipated that learner demand for data science education courses will increase if the KISTI operates online education courses that learners can complete in a short time to improve their capabilities during off-hours, such as time after work or weekends.

Furthermore, the KISTI should internally develop online education course operation strategies [7], such as stable server establishment, standardization of online education methods and operation, and an increase in instructor capabilities for providing online education, to successfully operate its online education courses. It is expected that these strategies can be proposed as new, alternative methods for the existing education system in circumstances of disasters such as the coronavirus pandemic of 2019 in which offline education courses are unlikely to be operational.

References

[1] Kwon, *The Sociology of Adult and Continued Education*. Hakjisa, 1999.

[2] P. Jarvis, *The Sociology of Adult and Continued Education*. Croom Helm, 1985.

[3] S. Merriam, *[title]*, [publisher], 2006.

[4] B. Trilling and C. Fadel, *21st Century Skills: Learning for Life in Our Times*. New York, NY, USA: Wiley & Sons, 2009.

[5] Han, "The Relationship between Academic Motivation and Self-Directed Leaning among Adult Learners" *The Journal of Learner-Centered Curriculum and Instruction*, Vol. 7(2), pp. 355-374, 2007.

[6] J. E. Lee, B. M. Yu and H. J. Park, "The Influence of the Presence Perceived by Learners and Participation Motivation on Satisfaction in Distance Education" *Journal of Agricultural Extension & Community Development*, Vol. 22(2), pp. 233-243, 2015.

[7] Lee, J. Dong, K. Misook, "University students' perceptions on the practices of online learning in the COVID-19 situation and future directions" *Multimedia-Assisted Language Learning*, Vol. 23(3), pp. 359-377, 2020.

Indexing processing method trend and performance analysis survey on Bio big data

Yejin Kwon, Jongsuk R. Lee, and Sunil Ahn
Department of National Institute of Supercomputing and Networking Center for Computational Science Platform, Korea Institute of Science and Technology Information (KISTI)
Daejeon, South Korea
[e-mail: yejinkwon@kisti.re.kr, jsruthlee@kisti.re.kr, siahn@kisti.re.kr]
*Corresponding author: Sunil Ahn

Abstract

The large-scale data processing system and indexing processing method are essential for bio big data including human genome information. Even if a database is constructed with only individual genome variant information as the subject, each individual has an average of 400 million to 500 million genome variant information. Furthermore, various big data processing technologies are required to process data of many people. Currently, genomic data is provided on a platform established by each nation medical research center for genomic data analysis worldwide.

Because those platforms are dealing with sensitive and highest level of security data, it has established and provided a process for data security and research ethics verification.

In this paper, we analyzed bio big data processing technology and research trends. Likewise, we analyzed the features of each genomic database. In addition, indexing processing performance for actual data was analyzed using data queries provided by each database.

***Keywords*:** bio bigdata, genomics, variant, nosql, database

1. Introduction

Bio big data refers to large amounts of data including human genomic information. Since GRCh38, the reference genome of humans, was completed in 2013, various methods for analyzing genome variant information based on individual genome maps are being studied[1]. In addition, as analysis tools, NGS(Next Generation Sequencing, that can quickly analyze genomes have been spread, the research environment has been established in which individuals' genome information can be accurately and quickly analyzed [2]. Therefore, various studies are being conducted based on the genomic information of each individual. GWAS(Genome-wide Association Study) studies are actively conducted to study the association with diseases by analyzing the genome mutation information possessed by each person [3].

Each individual has an average of 400 million to 500 million genome variations based on SNP(Single-nucleotide polymorphism). GWAS is a study that analyzes the relationship between the genome variations in human and a specific disease and derives the result [4]. The purpose of this study is to derive the relationship between

This research was supported by the Building and providing a research environment to utilize bio data Program through the National Research Foundation of Korea(NRF) (No. N-20-NM-CR12-S01), and the KISTI Program (No. K-20-L02-C05-S01)

the genetic variation and the disease by analyzing the genome variation of individuals with specific diseases [5]. In order to conduct such research, it is necessary to construct and indexing bio big data, and various processing modules for big data processing are essential to analyze the relationship between genome variations in human and diseases of each individual. At present, as various methodologies for building bio big data, big data processing and indexing methodology using nosql is the most widely used [11-16] . In this paper, we analyzed the case of service construction of an indexing processing system that constructs and searches human genomic variants information as a database, and analyzes the performance of the system to query a genome database.

2. The Trends of Genomic data Research

2.1 The Trends of Genomic data research by country

The major analysis of human genomics variant has been studied in Europe. In particular, Europe has been conducting research for more than 10 years based on the European Union.

For the study of genetic information on people analysis, because of the information is required highest level of security, the process is established so that all research projects for human genome analysis in each country are accessible only to researchers with specific access rights.

Table 1. Bio Bigdata Research platform by nation

Nation	Project Name	Site
USA	All of Us (2015)[6]	https://allofus.nih.gov/
Canada	Genomics in Canada (2000)[7]	https://www.genomecanada.ca/
Euro	genomics England (2013)[8] EuroBioBank (2013)[9]	https://www.genomicsengland.co.uk/ http://www.eurobiobank.org/
Japan	Japan Agency for Medical Researh and Development (2015)[10]	https://www.amed.go.jp

As described in **Table 1**, you can check the project's simple informaiton that start year and project name for human genome information for each major country. The construction of genomics database for all people does not allow general users to access, and is constructed so that genome information can be accessed through strict researcher screening procedures, security pledges, and researcher review committees.

Genomics England has analyzed the genome data of 100,000 people and is conducting analysis studies on rare diseases and cancers.

2.2 The trends of Genomics Database

Although access to human genome information is not allowed, the nosql database is mainly used as a service to build a database for human genome information and access to researchers.

Table 2. Genomics Database

Name	Database	Service
GEMINI[11]	SQLite	Framework
HGVA[12]	MongoDB Hbase	Genomics England
Genomics DB[13]	TileDB	Framework
GeneSysV[14]	ElasticSearch	Framework
RD-Connect[15]	Liferay	Biobank
GnomeAD[16]	GenomicsDB (Intel) ElasticSearch 5.5	Google Cloud

Table 2 summarizes how to store and indexing genome information of human large-scale big data, and what services are used.

GEMINI provides a framework that enables searching and indexing for each sample. Each researcher can install the GEMINI framework and use it for personal research. HGVA is built on the basis of MongoDB and Hbase. It is possible quick analysis of statistical information on genomic data. GenomicsDB can quickly build a database based on genomic information, but the search speed for each sample is relatively slow. In the case of GenSysV, a parent-child structure was designed to relatively increase the search speed for each sample. Therefore, the search speed for samples is fast, but the speed for updating each sample is slow. RD-Connect provides a service to define filtering and query for all information. In addition, GenomiAD is built on Apache Spark and boasts a very fast search speed, but does not support searching for all data. The analysis of the characteristics of this database is defined in Table 3 below.

Table 3. Bio big data processing method

Name	Advantage	Disadvantage
GEMINI	Sample Reference	Slow Search
HGVA	Query Aggregation	Sample Search
Genomics DB	Fast Importing	Sample Search
GeneSysV	Sample Search	Slow Importing
RD-Connect	Filtering Query Sample Catalog	Slow Search
GnomeAD	Fast Search	Data Coverage

3. Performance Comparison

Table 4 shows the results of response speed when searching for the search query CHROM: 6, POS (>=): 28477797, POS (<=): 33448354 based on the currently accessible database.

Table 4. Response Time on Database

Name	Query	Response Time
GEMINI	Select chrom, start, ref., alt, gene, max_aaf_all, impact, rs_ids from variants where chrom = 'chr6' and start > = 28,477,797 and end <= 33,448,354 and impact= 'missense_variant' limit 400	7.72s[14]
GeneSysV	CHROM: 6, POS (>=): 28477797, POS (<=): 33448354, Consequence: missense_variant.	6.77s[14]
RD-Connect	Filtering Query(Web UI)	5.86s
GnomeAD	6-28477797-33448354	0.042s

Not all conditions were measured in the same state, but when the same query was requested, the response time was measured for each platform or framework. The most recent GenomAD platform is built as a web-based system. GenomAD is built and indexed for data using Hail framework based on Apache Sparks, so it shows the fastest performance despite being a web-based platform.

4. Conclusions

In this paper, research trends and services for each country related to the construction of human genomic big data database were analyzed. In addition, the limitations of the services provided by each database and each feature were analyzed and summarized. Several services are provided that provide database of human genome variants information, but the system that builds and searches result data for a large number of human genetic information is currently in progress, and its features are clearly divided according to the database design method.

The database for storing genomic information of a large number of people is optimized for search speed or data importing function depending on the design method. In a framework that aims to convert new data into a database very quickly. It is meaningful to proceed with data construction rather than search speed. In

addition, the framework for quickly searching the established data also provides a service that optimizes the query in order to search at a high speed with the existing data. In order to construct such bio big data, technologies for processing various big data are being used, and differentiated technologies are being grafted according to the characteristics of each database.

Overseas, a system that processes and builds personal bio big data has been built and proceeded, but until now, systems for integrated genomic data management and processing in Korea are still insufficient. I hope that a platform service that integrates and manages each individual's genomic information and medical information and provides analysis tools will be conducted in Korea.

References

[1] Xiangqun Zheng-Bradley, Ian Streeter, Susan Fairley, David Richardson, Laura Clarke, Paul Flicek, the 1000 Genomes Project Consortium, "Alignment of 1000 Genomes Project reads to reference assembly GRCh38", GigaScience, Vol 6, Issue 7, July 2017, gix038, https://doi.org/10.1093/gigascience/gix038

[2] GU, Wei; MILLER, Steve; CHIU, Charles Y. "Clinical metagenomic next-generation sequencing for pathogen detection.", Annual Review of Pathology: Mechanisms of Disease vol. 14, 2019, pp. 319-338. https://doi.org/10.1146/annurev-pathmechdis-012418-012751

[3] Annalisa Buniello, et al, "The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019", Nucleic acids research vol. 47 Issue D1, pp. D1005–D1012, 2012. https://doi.org/10.1093/nar/gky1120

[4] WANG, Kai, et al. "PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data.", Genome research, , vol. 17.11, pp. 1665-1674, 2007. https://doi.org/10.1101/gr.6861907

[5] SYVÄNEN, Ann-Christine., "Toward genome-wide SNP genotyping" Nature genetics, vol. 37.6, pp. S5-S10, 2005. https://doi.org/10.1038/ng1558

[6] Sankar, Pamela L., and Lisa S. Parker., "The Precision Medicine Initiative's All of Us Research Program: an agenda for research on its ethical, legal, and social issues." Genetics in Medicine vol. 19.7 pp. 743-750, 2017. https://doi.org/10.1038/gim.2016.183

[7] Lemay, Margaret A. "The Role of Expectations of Science in Shaping Research Policy: A Discursive Analysis of the Creation of Genome Canada." Minerva vol 58, pp. 1-26, 2020. https://doi.org/10.1007/s11024-020-09395-5

[8] Samuel, Gabrielle Natalie, and Bobbie Farsides., "Public trust and 'ethics review'as a commodity: the case of Genomics England Limited and the UK's 100,000 genomes project." Medicine, Health Care and Philosophy vol. 21.2, pp. 159-168, 2018. https://doi.org/10.1007/s11019-017-9810-1

[9] Mora, Marina, et al., "The EuroBioBank Network: 10 years of hands-on experience of collaborative, transnational biobanking for rare diseases." European Journal of Human Genetics vol. 23.9, pp. 1116-1123, 2015. https://doi.org/10.1038/ejhg.2014.272

[10] Suematsu, Makoto. "Japan Agency for Medical Research and Development (AMED)." Science Impact vol 11.4, pp. 48-49, 2017. https://doi.org/10.21820/23987073.2017.4.48

[11] PAILA, Umadevi, et al., "GEMINI: integrative exploration of genetic variation and genome annotations.", PLoS Comput Biol vol. 9.7, e1003153, 2013. https://doi.org/10.1371/journal.pcbi.1003153

[12] Lopez, Javier, et al., "HGVA: the Human Genome Variation Archive.", Nucleic acids research vol. 45. Issue W1, pp.: W189-W194, 2017. https://doi.org/10.1093/nar/gkx445

[13] Datta, K., et al. "GenomicsDB: storing genome data as sparse columnar arrays." White Paper. Intel Health and Life Sciences, 2017.

[14] Zia, Mohammad, et al., "GenESysV: a fast, intuitive and scalable genome exploration

open source tool for variants generated from high-throughput sequencing projects.", BMC bioinformatics vol. 20.1, pp. 61, 2019. https://doi.org/10.1186/s12859-019-2636-5

[15] Thompson, Rachel, et al., "RD-Connect: an integrated platform connecting databases, registries, biobanks and clinical bioinformatics for rare disease research.", Journal of general internal medicine vol. 29.3, pp. 780-787, 2014. https://doi.org/10.1007/s11606-014-2908-8

[16] Koch, Linda. "Exploring human genomic diversity with gnomAD." Nature Reviews Genetics vol. 21.8, pp. 448-448, 2020. https://doi.org/10.1038/s41576-020-0255-7

Design of GUI-based deep learning neural network service model to support various frameworks

Jae-Gwang Lee, Jongsuk R. Lee, and Sunil Ahn
Department of National Institute of Supercomputing and Networking Center for Computational Science Platform, Korea Institute of Science and Technology Information (KISTI)
Daejeon, South Korea
[e-mail: leejg@ kisti.re.kr, jsruthlee@kisti.re.kr, siahn@kisti.re.kr]
*Corresponding author: Sunil Ahn

Abstract

Various frameworks exist for designing deep learning neural networks. Recently, a GUI-based deep learning neural network design tool has been developed to easily design a deep learning neural network. However, since the existing GUI-based deep learning neural network design tool does not support various frameworks, it is impossible to use different tools for each framework or trained models in different frameworks. In this paper, the open source ONNX standard is used to support various frameworks. Since ONNX standard cannot be expressed with GUI, it converts JSON format to and applies GUI. The applied GUI-based deep learning neural network design has the advantage that it is faster and easier than developing a code-type deep learning neural network in block form. In addition, in order to cope with various frameworks with rapid changes, the ONNX standard-based schema was developed and designed to easily add and modify ONNX operators. Therefore, in this paper, we propose a service model that can modify and design deep learning neural network models of various frameworks based on GUI using ONNX standard and JSON format.

***Keywords*:** GUI, Deep learning, Nerual network Design, ONNX

1. Introduction

Machine learning is a field of artificial intelligence, which is a technique in which a computer learns data and makes judgments or predictions[1]. Deep learning is a field of machine learning. It is a technology that predicts and judges through a large amount of computation by constructing a network similar to the structure of the brain. As the performance of hardware improves due to the development of IT technology, it is recently applied to various fields, and research is actively being conducted[2].

As research on deep learning became active, various open source frameworks for deep learning development were published[3].

In the case of framework, many improvements are being made through updates. However, the framework used is different depending on the data and what you are trying to predict. In other words, there is a regret that the framework used is different depending on the field and occupation group such as researchers, developers, etc. This problem has the disadvantage of having to learn another framework in order to use a model designed in another framework[4].

This research was supported by the KISTI Program (No. K-20-L02-C05-S01), the EDISON Program through the National Research Foundation of Korea (NRF) (No. NRF-2011-0020576). A Grant was awarded by the Ministry of Science and ICT (MSIT) under the Program for Returners for R&D.

In addition, it is difficult for beginners to approach deep learning neural network design. The reason is that in order to design a deep learning neural network, you need to learn a new language. In order to solve these problems, research and development on GUI(graphical user interface)-based deep learning neural networks are being conducted[5].
However, the existing GUI-based deep learning neural network tool has a disadvantage in that it is difficult for beginners who do not know deep learning to access it because it is complex, or it is difficult to extend because it is dependent on a specific framework.
In this paper, we propose a system that allows anyone to quickly and easily design deep learning neural networks using drag & drop method and algorithm template method.

2. Related research

2.1 Expresso

Expresso is A User-Friendly GUI for Designing, Training and Exploring Convolutional Neural Networks. Expresso is open source used to develop CNN(Convolutional Neural Network).

Expresso provides a convenient wizard-like graphical interface that guides users through a variety of common scenarios such as importing data, building and training deep networks, performing various experiments, analyzing and visualizing these experimental results[7].
Because Expresso uses a specific framework, it cannot interact with other frameworks, and has the disadvantage of configuring only CNNs.

2.2 DL-Dashboard

DL-Dashboard is User-friendly Deep Learning Model Development Environment . DL-Dashboard supports a deep learning model development environment of drag&drop method. To design a deep learning model, it is impossible to understand a specific framework and apply the model designed to other frameworks. It takes a framework-independent approach. It provides components such as convolutional layer, pooling layer, and batch normalization, and can be used by configuring CNN and RNN[6]. However, DL-Dashboard takes a method that can be easily used by users who understand deep learning.

2.3 ONNX

ONNX(Open Neural Network Exchange) is an open source format for constructing machine learning models and supports various frameworks. ONNX defines built-in operators, standard data types[8]. ONNX defines built-in operators, standard data types. In addition, the file format is defined so that the model can be used with various frameworks, tools, runtimes and compilers. This simplifies the conversion between different frameworks and increases usability[9].
In this paper, a GUI-based deep learning neural network design system using features of ONNX's(Open Neural Network Exchange) various framework support was proposed.

3. Manipulation of Virtual Model

3.1 Requirements Analysis

In this paper, we propose a GUI-based deep learning neural network design service model. The requirements for the service model are as follows.
· Usability: A dreg and drop type deep learning neural network design tool that can be easily used by users is required. In addition, it should be possible to easily construct a deep learning neural network.
· Scalability: As the version changes rapidly such as framework and onnx, it is necessary to support rapid response to changes based on the operator of the onnx standard.
· Dependency: Deep learning neural network models designed in other frameworks can also be imported, and models can be modified and reused in a GUI environment.
In this paper, a service model is proposed so that it is easy to use, can modify the model designed in various frameworks, and can respond quickly to changes.

3.2 Proposed service model

Fig. 1 is a GUI-based deep learning neural network design service model proposed in this paper.
The service model proposed in this paper can be divided into two categories: JSON format definition for import/export and GUI deep learning neural network design.

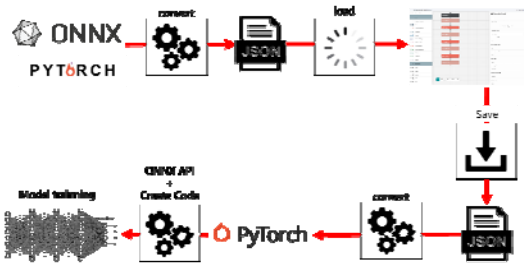


Fig. 1. Proposed service model

3.3 JSON schema

The JSON format consists of a node to design a model, a layer group to manage the node, a template to quickly and easily construct a deep learning neural network, and a JSON to manage the designed deep learning neural network.

The json format that composes the first node was defined by analyzing pytorch's nn library and ONNX operator. node has parameter name, default value, location value, etc.

The second layer group is defined to facilitate the addition and deletion of nodes by grouping nodes. The layer group is composed of a total of 13 types, such as convolutional layer, pooling layer, and normalization layer.

The third template json format is defined so that it can be modified by connecting specific nodes to form a deep learning neural network.

Finally, deep learning neural network json is used to save and exprot neural networks designed in GUI-based deep learning neural network tools, or to import neural networks designed in other frameworks.

The import module converts neural networks designed in various frameworks into ONNX expressions, and then converts ONNX expressions into defined JSON format, which can be modified in GUI tools.

The export module can be saved as '.py' that composes Pytorxh's neural network in the defined JSON format, or it can be saved as JSON, and it can be linked with AI Stduio to perform training based on HPC.

3.4 GUI-based deep learning neural network design

The deep learning neural network configuration tool consists of component, designer, and configuration.

Component consists of node, template, and save_model, and node is configured by analyzing ONNX operator and nn of pytorch. Using nodes, it is possible to construct a deep learning neural network desired by the user. In addition, when using a template, a neural network is automatically configured and the configured neural network can be modified by dragging and dropping the desired algorithm on the designer. In addition, you can modularize your own neural network to the save model with your own neural network.

The designer can configure a neural network by dragging and dropping nodes. Connect nodes to form a network.

In the configuration, parameter values of each node can be entered and modified, and the node operator type can be changed. First, GUI Deep Learning neural network design can be modified and reused after designing a neural network using the dreg and drop method or importing an existing model. **Fig. 2** shows the developed GUI-based deep learning neural network design tool.

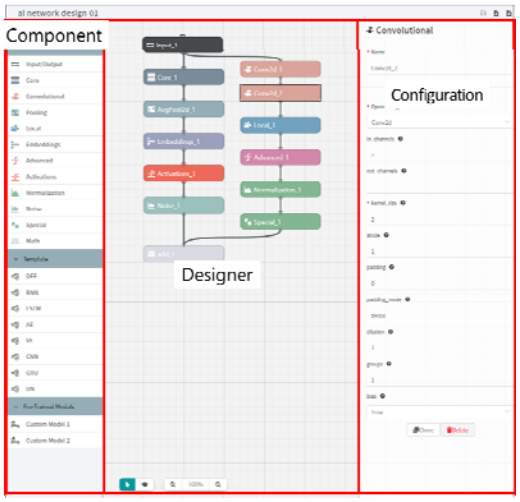


Fig. 2. GUI-based deep learning neural network design tool

4. Conclusions

In this paper, we designed and developed a GUI tool that allows users to easily and simply design deep learning neural networks. In addition, the concept of template was introduced so that a deep learning neural network can be designed without knowing the framework. It is expected that even beginners can easily design deep learning neural networks with templates.

In addition, in order to avoid being dependent on one framework, ONNX was applied so that models designed in other deep learning frameworks can also be imported.
A simple model test was performed using MNIST data. It is expected to be easy and simple to use for various users (beginner to advanced). For future research, we will conduct research that applies AutoML and Generative AI.

References

[1] Jinho Jeong, “Deep Learning Framework Trend and Introduction”, ICT Standard Weekly, vol.968, pp.1-4, 2020
[2] Y. LeCun, Y. Bengio and G. Hinton, "Deep learning," Nature, vol. 521, (7553), pp. 436-444, 2015
[3] Jihun Chae, Hyoung-yong Ko, Byoung-gil Lee, Namgi Kim, “A Study on the Pipe Position Estimation in GPR Images Using Deep Learning Based Convolutional Neural Network" Journal of Internet Computing and Services, Vol.20, No.4, pp.39-46, 2019.
[4] Yo-Seob Lee, Phil-Joo Moon, “A Comparison and Analysis of Deep Learning Framework”, Journal of the KIECS. Vol.12. No.1, pp.115-122, 2017
[5] Aditya Akundi, Tzu-Liang Bill Tseng, Zejing Cao, Hoejin Kim, “A Deep Learning Graphical User Interface Application on MATLAB”, 2018 ASEE Annual Conference & Exposition, 2018
[6] Yoomi Park, Eun-Ji Lim, Shinyoung Ahn, Wan Choi, Taewoo Kim, “DL-Dashboard: User-friendly Deep Learning Model Development Environment”, IUI '19: Proceedings of the 24th International Conference on Intelligent User Interfaces: Companion, pp.73-74, 2019
[7] Ravi Kiran Sarvadevabhatla, R. Venkatesh Babu, “Expresso: A User-Friendly GUI for Designing, Training and Exploring Convolutional Neural Networks”, arXiv:1505.06605v2 [cs.CV] 18, 2015
[8] Xinli Cai, Peng Zhou, Shuhan Ding, Guoyang Chen, Weifeng Zhang, ”Sionnx: Automatic Unit Test Generator for ONNX Conformance“, arXiv:1906.05676v1 [cs.SE], 2019
[9] Open Neural Network Exchange, https://onnx.ai/

KSII The 12th International Conference on Internet (ICONI) 2020.

Design of Simulation Job Cache System on a Computational Science Platform

Inho Jeon, Noori On, Nam-Gyu Kim and Jong Suk Ruth Lee
Div. of National Supercomputing, Korea Institute of Science and Technology Information (KISTI)
Deajeon, South Korea
[e-mail: {inojeon, claireon, ssgye, jsruthlee}@kisti.re.kr]
*Corresponding author: Jong Suk Ruth Lee

Abstract

In this paper, we propose a Simulation Job Cache system that provides an environment in which more users can effectively use computing resources by minimizing unnecessary processes in the Computational Science Platform (CSP). The goal of the proposed system is to provide an environment in which more users can effectively utilize computing resources by minimizing unnecessary processes of CSP. Compared to the existing system that executes all the tasks requested by the user, the Computational Science software, which has many repetitive executions, efficiently utilizes resources and provides users with quick execution results

***Keywords*:** Computational Science Platform, Cache system, Web-based Simulation

1. Introduction

Due to the influence of COVID-19, most of the science and engineering college classes are being conducted non-face-to-face. Theoretical classes were able to proceed with lecture-based and discussion-based classes by utilizing e-Learning contents and video conferencing solutions such as Zoom[1]. However, in the case of practical classes, it is very difficult to proceed to the level before Corona 19 due to space constraints. An online virtual lab environment is in the spotlight as one of the alternatives to solving these problems. In the future, the ability to use computers to manipulate data related to one's major will be more important than dealing with old laboratory equipment in school labs. In order to provide a non-face-to-face practice environment, KISTI is servicing the Computational Science Platform (CSP), an online environment in which Computational Science(CS) software can be executed. It is possible to easily register software developed by multi-disciplinary computational science researchers and build an environment that other users can run. Platform users can run the computational science software they want through a web connection without the installation process[2-3].

Users who use the CSP in practice classes have the following characteristics. First, in the practice class, users use the same software to solve the same problem, and execute the software using the same input files. Second, users run simulation software during the class project period. When a large number of users use the service at the same time, a problem arises that the service is slow or the waiting time for submitting a job is long due to overload of the service and the lack of available computational resources. The system needs to be improved to cope with this problem.

In this paper, we propose a Simulation Job Cache system that provides an environment in which more users can effectively use computing resources by minimizing unnecessary processes in the CSP. Based on the Simulation Job data

This research was supported by the EDISON Program through the National Research Foundation of Korea(NRF) (No. NRF-2011-0020576), and the KISTI Program (No. K-20-L02-C05-S01)

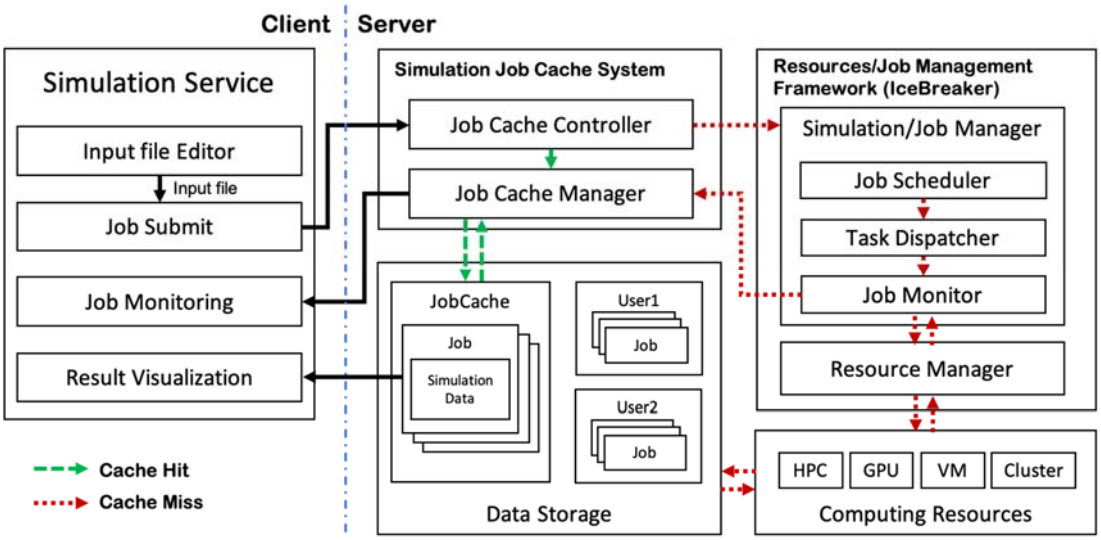


Fig. 2. Architecture of the Simulation Job Cache System

used in the existing practical classes, it was confirmed that the proposed system showed superior performance compared to the existing system.

2. Related Research

2.1 Caching

Caching speeds up application processing. Storing a copy of the already fetched data or computed results speeds up processing, which makes it possible to process future requests faster. Caching is an efficient architectural pattern because most of the same data programming accesses instructions repeatedly[4]. **Fig. 1** shows a diagram of the cache system on a simple Web server that provides static resources upon request from a web browser. In the web server, if there is no cache for the browser's request, a cache miss occurs and data is requested from the hard disk. Then, it returns the requested data and stores the data in the cache. Subsequently, a cache hit is generated for the same request and the data stored in the cache is returned.

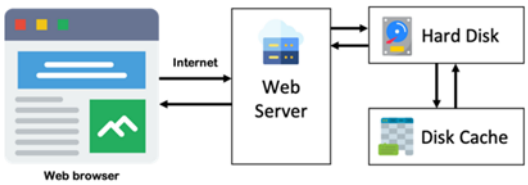


Fig. 1. Diagram of the cache system on a simple Web server

3. Simulation Job Cache System

The process of using CS software for education is as follows. First, introduce the theory in the subject, then introduce the practical method using CS software. Students change and execute input variables of CS software, and solve tasks by analyzing execution results. In order to solve the systemic problems occurring in this process, we propose a simulation job cache system.

Fig. 2 is a schematic diagram of the proposed system. The job is not executed under the user account requested by the client, but the job is executed through a common JobCache account. The proposed system consists of three main modules.

- Simulation Job Cache Storage: A data store where unique job data is cached.
- Job Cache Controller: Determine whether to execute CS software for the job requested to the user or to provide cached data.
- Job Cache Manager: Manages data in cache storage and sends data related to jobs to clients.

In the existing CSP, the execution time T_{total} taken when executing N tasks for one software

Table 1. Comparative Analysis of Performance by CS software (A: existing system, B: proposed system)

CS Software	Total job counts	Unique job counts	Duplicate percent	Avg. run time(s)		Total Storage(MB)		Total resource usage time(h)	
				A	B	A	B	A	B
roundSTMtip	2,684	359	87 %	54	8.2	614.6	82.2	40	5
acuteSTMtip	2,424	374	85 %	55	9.5	555.1	85.6	37	6
pianostring	3,595	203	94 %	972	55.9	21,537.6	1,216.2	971	55
gravityslingshot	7,028	5,415	23 %	1,425	1,102.9	47,790.4	36,822.0	2,782	2,143

can be expressed as follows.

$$T_{total} = \sum_{n=0}^{N-1} t_{jn}$$

Where, t_{jn} represents the time taken when the nth task is executed. When the proposed system is applied, the time $T_{perpose}$ takes to execute N tasks is as follows.

$$T_{perpose} = \sum_{n=0}^{U-1} t_{jn} + \sum_{n=0}^{N-1} t_{cn}$$

Where, U is the number of unique jobs that are not duplicated among N jobs, less than or equal to N. t_{cn} is the time taken to determine whether the n-th job is a cache hit in the job cache controller. In general, t_{jn} is greater than t_{cn}. When users solve the same problem, the value of U becomes sufficiently smaller than the value of N, and improved performance compared to the existing system.

Table 1 is a table that compares and analyzes the performance of the existing system and the proposed system for the CS software used in the practice class.

4. Conclusion and Future Work

In this paper, we proposed a simulation job cache system that minimizes the execution of duplicated job in order to efficiently use computing resources in the context of using the CSP for education. Compared to the existing system that executes all the tasks requested by the user, the CS software, which has many repetitive executions, efficiently utilizes resources and provides users with quick execution results. However, if it is not for educational purposes, the same job is rarely submitted. In this case, when the proposed system is applied, system performance may be degraded. In the platform, it is necessary to distinguish between educational users and research users. If the proposed system is provided to educational users, a better simulation execution environment can be provided.

In the future, we plan to implement the designed system. We plan to develop an API of the same type as the restful API provided by Icebreaker, and this can minimize client code modification. Comparison with other systems designed for similar purposes will demonstrate the superiority of the proposed system.

References

[1] Sang-Mi Kim (2020), Analysis of Press Articles in Korean Media on Online Education related to COVID-19, *Journal of Digital Contents Society*, Vol. 21, No. 6, pp. 1091-1100, DOI: 10.9728/dcs.2020.21.6.1091

[2] Inho Jeon, Yejin Kwon, Jin Ma, Sik Lee, Kum Won Cho, & Jerry Seo (2017). Construction and Service of a Web-based Simulation software management system for the Computational Science and Engineering. *Journal of Internet Computing and Services*, 18(4), 99-108. DOI: 10.7472/jksii.2017.18.4.99.

[3] Yejin Kwon, Inho Jeon, Jerry H. Seo, & Jongsuk R. Lee (2020). Interface Development for Pre and Post processor on EDISON Platform Simulation System. *Journal of Internet Computing and Services*, 21(1), 103-110. DOI: 10.7472/jksii.2020.21.1.103.

[4] Nick Karnik, "The hidden components of Web Caching", Feb, 2018, (https://www.freecodecamp.org/news/the-hidden-components-of-web-caching-970854fe2c49/)

Pancreas Segmentation using Deep Learning Model

Sang-Yun Lee, Byoung-Dai Lee and Namgi Kim
School of Computer Science and Engineering, Kyounggi University
Suwon, South Korea
[e-mail: e311333111313@gmail.com, {blee, ngkim}@kgu.ac.kr]
*Corresponding author: Namgi Kim

Abstract

The pancreas, which is quite small and overlaps with other internal organs, is difficult to segment by accurately finding its region in computed tomography (CT) images. This paper proposes a multiple-concatenated U-Net (MCU-Net) model that can more accurately segment the pancreas region in a CT image through machine learning. The MCU-Net model concatenates several U-Nets with different filter sizes on the final layer. The experimental results show that the accuracy of the MCU-Net model is higher than that of conventional U-Net models.

***Keywords*:** Pancreas segmentation, Deep learning model, Artifical intelligence, Image processing

1. Introduction

Recently, computer-aided detection systems have been actively emerging. They use convolutional neural network (CNN)–based deep-learning technology to segment the region of a particular organ or identify its volume in computed tomography (CT) images. However, even if such technology is used, accurately segmenting the pancreas region is more difficult than other organs. Therefore, this paper proposes a deep-learning model that can improve the accuracy of pancreas-region segmentation.

2. Related Work

The U-Net [1] model, which segments a specific organ or cell from the input image, is composed of a U-shaped network. The U-Net excludes minor image parts through a contracting path and an expanding path, and delivers the context information that becomes a feature to the end of the network. The U-Net uses reflection padding that places values opposite an edge to capture the context information of the edge without loss, when passing through the filter. Therefore, it loses less information than the zero-padding method, which puts a zero in the edge value.

The attention U-Net [2] model, a modified version of the U-Net model, adds an attention gate that can focus on the features. The attention gate helps to selectively maintain the image-feature information. The models trained using the attention gate mainly learn the distinctive features of an image to better predict the target. They simultaneously and implicitly learn how to suppress the regions not related to the target. This has improved the models' sensitivity and accuracy.

The U-Net and the attention U-Net models have a rapid training speed. However, after training the model using the data, it is observed that the

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2020R1A6A1A03040583 and NRF-2017R1D1A1B04027874).

time required for training is not significant in a deep-learning–based automated segmentation system in which only the target image is input into the trained model. Therefore, this paper proposes an MCU-Net model that concatenates several U-Nets to more accurately segment the pancreas region.

3. Proposed Method

The MCU-Net model, an extended version of U-Net, extracts features by traversing the input image in patch units. The MCU-Net model can obtain more context information than the U-Net model using mutiple patch sizes. **Fig. 1** shows the structure of the MCU-Net model, which has three concatenated U-Nets with patch sizes (3, 3), (5, 5), and (7, 7). In the figure, the number above each layer represents the number of channels, and the lower-left number represents the size of the feature map for the x-y size. The gray layers indicate that the feature maps have been copied and concatenated.

4. Experiments

In the experiments, model training and evaluation were performed using the Pancreas-CT set [3], which is the most widely used for pancreas segmentation. In this study, the proposed MCU-Net model was implemented, based on the source code provided in the U-Net-based model [4]. Python version 2.7.12 was used for the experiments, and the model was implemented using the TensorFlow (ver. 2.1.0) and Keras (ver. 2.0.8) Python-based deep-learning libraries. The experiment was performed using Ubuntu 16.04 LTS on a processor with 16 GB RAM and a GeForce GTX 2080 Ti graphic card. Additional libraries included NumPy (ver. 1.14 or higher), pandas, and matplotlib. For the hyper-parameters used in the experiment, 10 was used for the epoch, 1e-5 for the learning rate, and Adam for the optimization function.

Table. 1. Comparison of the MCU-Net and U-Net models

Dataset (fold)	1		2		3		4		Total
	AVG	DEV	AVG	DEV	AVG	DEV	AVG	DEV	AVG
M1	0.864	0.038	0.906	0.013	0.907	0.010	0.915	0.014	0.896
M2	0.868	0.034	0.917	0.009	0.918	0.008	0.923	0.012	0.906
M3	0.860	0.034	0.918	0.009	0.917	0.008	0.923	0.011	0.904
M4	0.862	0.040	0.915	0.011	0.917	0.008	0.921	0.011	0.903
M5	**0.876**	0.033	**0.928**	0.008	**0.929**	0.008	**0.931**	0.008	**0.916**

Table 1 shows the dice similarity coefficient (DSC) results of this experiment. These are the average results of when the corresponding fold was used for the test results. In Table 1, model M1 represents a U-Net model [5] wherein a conventionally designed U-Net–based baseline model [1] is combined with the pancreas classification. Model M2 is an ensemble model, in which a softmax was executed by simply adding two U-Nets of the same patch size (3, 3). Model M3 has ensembled three U-Nets, each

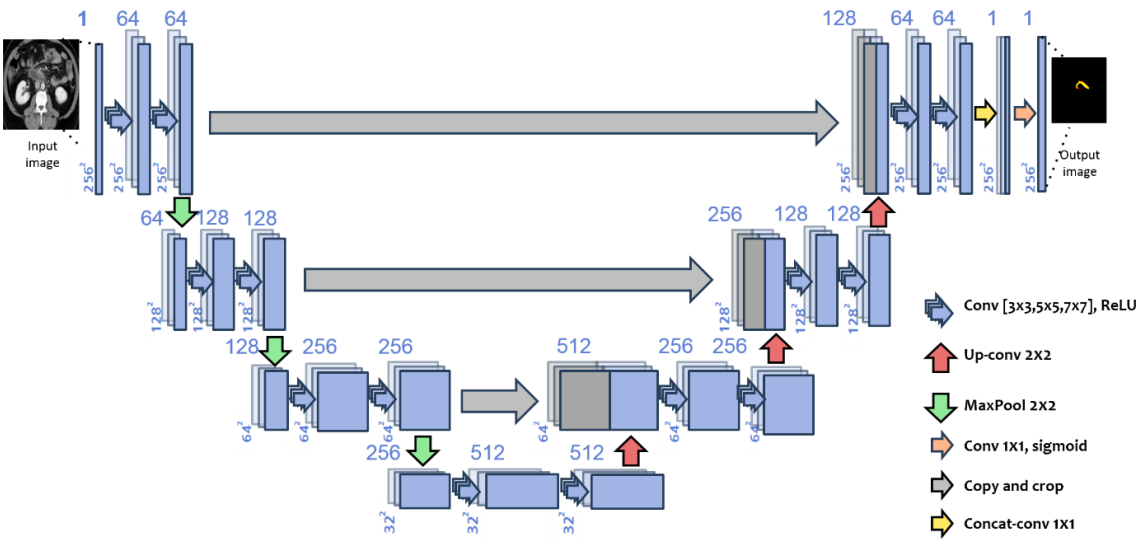


Fig. 1. Proposed model using three patch sizes (3×3, 5×5, and 7×7)

with a patch size of (3, 3). Model M4 is a multi U-Net model consisting of two U-Nets with patch sizes of (3, 3) and (5, 5). Model M5 is a multi U-Net model consisting of three U-Nets with patch sizes of (3, 3), (5, 5), and (7, 7).

5. Conclusions

This paper proposed the MCU-Net model and it showed the 91.6% accuracy, which was 1.8% higher than that of the U-Net–based model [5]. In a future study, we plan to investigate whether significant results will be produced when other CT datasets are used. In addition, we will verify whether this model produces significant results in areas where CNNs that operate on images can be applied. Afterwards, multiple concatenations will be applied to models other than U-Net to investigate whether the performance improves.

References

[1] O. Ronneberger, P. Fischer, T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015.

[2] O. Oktay, et al., "Attention u-net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999*, 2018.

[3] H. Roth, et al., “Data from pancreas-CT. The Cancer Imaging Archive,” http://doi.org/10.7937/K9/TCIA.2016.tNB1kqBU, 2016.

[4] https://github.com/snapfinger/pancreas-seg

[5] Y. Liu, S. Liu, “U-NET for pancreas segmentation in abdominal CT scans,” *ISBI Challenge*, 2018.

An investigation on the effect of imputing missing values on time series forecasting model performance

Hyun Ahn[1,2], Dinh-Lam Pham[1,2], Kyoung-Sook Kim[2], Haesung Lee[3], Minjae Park[4] and Kwanghoon Pio Kim[1,2*]
[1]Division of Computer Science and Engineering, Kyonggi University
Gyeonggi, South Korea
[2]Contents Convergence Software Research Institute, Kyonggi University
Gyeonggi, South Korea
[3]Smart Power Distribution Laboratory, KEPCO Research Institute
Daejeon, South Korea
[4]Division of Computer Information, Daelim University College
Gyeonggi, South Korea
[e-mail: {hahnkgu, dinhlampham, khmjmc, kwang}@kyonggi.ac.kr, seastar@kepco.co.kr, mjpark@daelim.ac.kr]
*Corresponding author: Kwanghoon Pio Kim

Abstract

Time series forecasting has a slew of applications for solving many real-world problems. However, we often come across undesirable missing values and these might be problematic to a forecasting task. In this paper, we investigate on the effect of imputing missing values on time series forecasting. As an instance of forecasting problem, we address the daily peak load forecasting at EV (electric vehicle) charging stations in South Korea. First, we train forecasting models based on LSTM (Long Short-Term Memory) with data imputation techniques. Then, we compare prediction accuracy of forecasting models so that we recognize an impact of missing value imputations compared with the no-imputation approach.

***Keywords*:** Time series forecasting, missing values, data imputation

1. Introduction

Time series forecasting has a slew of applications for solving many real-world problems from power system [1] to business process prediction [2]. However, we often come across undesirable missing values and these might be problematic to a forecasting task because they can cause a wrong bias or reduced volume of data.

In this paper, we investigate on the effect of imputing missing values on time series forecasting. As an instance of forecasting problem, we address the daily peak load forecasting at EV (electric vehicle) charging stations in South Korea. First, we train forecasting models based on LSTM (Long Short-Term Memory [3]) deep learning model with data imputation techniques. Then, we compare prediction accuracy of forecasting models so that we recognize an impact of missing value imputations compared with the no-imputation approach.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (grant number 2020R1A6A1A03040583).

2. Forecasting Problem

As an instance of time series forecasting problem, we address the daily peak load forecasting of EV charging stations. Each EV charging station has different charging patterns depending of the type of major users and spatial characteristics (e.g., vehicular flow). Load forecasting allows us to understand the usage patterns of individual charging stations. **Fig. 1** shows the measurement of the daily peak load that is the target variable of the forecasting model. Each charging station usually operates multiple chargers, and the dataset consists of charging records collected from the individual chargers.

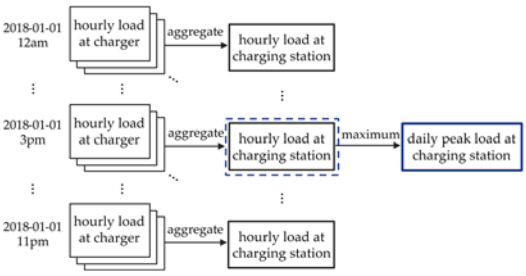


Fig. 1. Measurement of the peak load at an EV charging station

In order to make a training set from the dataset, we designed a feature set containing the following seven features:

- Charging station ID
- Day of the week
- Num. of normal/fast charges
- Amounts of normal/fast charges
- Daily peak load

3. Experimental Results

To investigate the effect of missing data imputation, we conduct an experiment according to the following: (1) a group of training sets is prepared by applying different imputation technique. (2) we obtain LSTM models derived from the training sets, each of which corresponds to each imputation technique. (3) we measure the prediction accuracy of each LSTM model and compare them to evaluate the effectiveness of imputation. For the experiment, we made use of a desktop computer powered by a Nvidia GeForce 2080 RTX Super and the PyTorch deep learning framework with data imputation libraries.

3.1 Dataset

Table 1 represents a summary of the dataset provided by KEPCO (Korea Electric Power Corporation). This dataset includes the load charging data collected from the EV charging stations in South Korea. In order to investigate the effect of missing rate on imputation performance, we prepare a total of five data sets by the threshold of missing rate (θ).

Table 1. Summary of the dataset

θ	Num. of stations	Data size	Num. of missing values	Missing rate
0.1	52	19,660	853	4.34%
0.2	96	37,820	3,705	9.80%
0.3	155	65,591	10,684	16.29%
0.4	223	96,204	21,218	22.06%
0.5	303	135,207	38,799	28.70%

3.2 Imputation results

We exploit the following six imputation techniques: LOCF (Last Observation Carried Forward), NOCB (Next Observation Carried Backward), EM (Expectation-Maximization [4]), k-NN (k-Nearest Neighbors), QUAD (quadratic spline interpolation) and QEM (quadratic spline interpolation with EM [5]).

Fig. 1 shows the distributions of target values obtained by performing imputation on the training set ($\theta = 0.5$). LOCF, NOCB, and k-NN generates substitute target values that approximate the population distribution. On the contrary, EM produces target values that underfitting the population distribution, which may lead to a less accurate model.

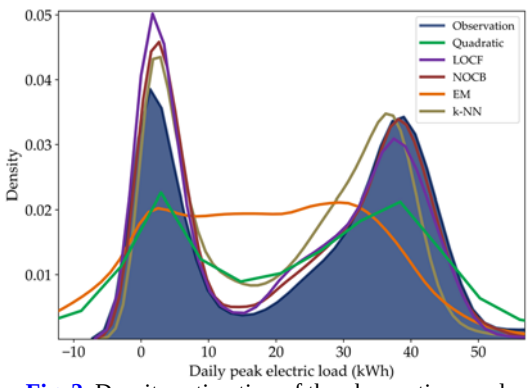


Fig. 2. Density estimation of the observations and imputed values

The prediction accuracy is measured by the mean absolute error (MAE) and MAE difference (δ) between a forecasting model and the no-imputation model (NI). **Fig. 3** shows the comparison of δ with the training set with $\theta = 0.5$. As the missing rate increases, the imputation models output a positive δ in many cases, and the magnitude of δ also surges.

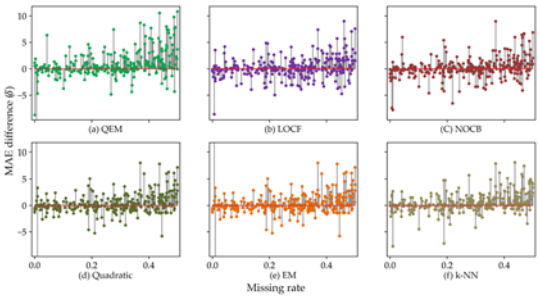


Fig. 3. MAE differences between each model and NI

A summary of the prediction accuracy with data imputation is presented in **Table 2**. In the case of QEM, this approach achieved the finest results in terms of prediction accuracy in three cases ($\theta = 0.2, 0.4$, and 0.5). Compared to NI, QEM reduced MAE by up to 9.8% at $\theta = 0.1$. In the results of EM and k-NN as multivariate imputation models, each of them showed the best accuracy in one case.

4. Conclusions

In this paper, we investigate the effect of imputing missing values at the time series forecasting problem. For this end, we constructed a peak load forecasting at EV charging stations based on the LSTM model. The experimental results showed that QEM has the best improvements in prediction accuracy for most of the cases. In the future, we are planning to devise an imputation method to obtain a reliable forecasting model that works well regardless of missing rates.

References

[1] W. Kong, et al., "Short-term residential load forecasting based on LSTM recurrent neural network," *IEEE Transactions on Smart Grid*, Vol. 10, pp. 841-851, 2017.

[2] S. H. Ham, H. Ahn, K. P. Kim, "LSTM-based business process remaining time prediction model featured in activity-centric normalization techniques," *Journal of Internet Computing and Services*, Vol. 21, No. 3, pp. 83-92, 2020.

[3] S. Hochreiter, J. Schmidhuber, "Long short-term memory," *Neural Computing*, Vol. 9, pp. 1735-1780, 1997.

[4] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Magazine*, Vol. 13, No. 6, pp. 47-60, 1996.

[5] B. Lee, H. Lee, H. Ahn, "Improving load forecasting of electric vehicle charging stations through missing data imputation," *Energies*, Vol. 13, No. 18, pp. 4893, 2020.

Table 2. Summary of the imputation results

θ		NI	QEM	LOCF	NOCB	QUAD	EM	k-NN
0.1	MAE	11.371 ± 0.191	11.448 ± 0.012	11.936 ± 0.003	11.522 ± 0.187	11.336 ± 0.159	11.306 ± 0.046	**11.268 ± 0.001**
	δ	–	-0.077	-0.565	-0.151	0.035	0.065	**0.103**
0.2	MAE	10.227 ± 0.03	**10.076 ± 0.036**	10.228 ± 0.044	10.174 ± 0.056	10.284 ± 0.029	10.213 ± 0.099	10.509 ± 0.026
	δ	–	**0.151**	-0.001	0.053	-0.057	0.014	-0.282
0.3	MAE	9.798 ± 0.007	9.681 ± 0.008	9.701 ± 0.022	9.726 ± 0.046	9.802 ± 0.047	**9.587 ± 0.057**	9.63 ± 0.004
	δ	–	0.117	0.097	0.072	-0.004	**0.211**	0.168
0.4	MAE	9.576 ± 0.031	**9.229 ± 0.05**	9.363 ± 0.085	9.515 ± 0.106	9.487 ± 0.088	9.278 ± 0.039	9.281 ± 0.067
	δ	–	**0.347**	0.213	0.061	0.089	0.298	0.295
0.5	MAE	9.47 ± 0.057	**8.537 ± 0.076**	9.192 ± 0.04	9.277 ± 0.038	9.122 ± 0.027	8.944 ± 0.135	8.982 ± 0.005
	δ	–	**0.933**	0.278	0.193	0.348	0.526	0.488

CNN Internal Operation Analysis Model using Feature Map

Ji-Soo Kang[1] and Ji-Won Baek[1] and Hyun Yoo[2]
[1]Department of Computer Science, Kyonggi University,
Gyeonggi-do, South Korea
[2]Contents Convergence Software Research Center, Kyonggi University
Suwon-si, South Korea
[e-mail: kangjs920@gmail.com, jwbaek@kyonggi.ac.kr, rhpa0916@gmail.com]
*Corresponding author: Hyun Yoo

Abstract

This paper presents a method of analyzing the inside of the artificial neural network composed of CNN using Feature Map. The latest deep learning algorithm is on the trend of large-scale, making it difficult to analyze within the neural network. The proposed method constructs a CNN-based image analysis model and detects changes in results resulting from partial changes in input data for the configured model. This enables visual analysis of the importance of each location of the image data. For this purpose, an artificial neural network is configured with ResNet and each pixel of the image that needs to be analyzed is reversed sequentially. In this process, positive and negative changes in the value of the node with the maximum neural output node are stored in the Feature Map. The two created Feature Maps are visually expressable. As a result, behavioral characteristics within the artificial neural network are expressed as input data values change. This method is visually expressable, so it helps a lot to interpret the results of deep neural networks.

Keywords: eXplainable Artificial Intelligence, Convolutional Neural Network, Deep Learning, ResNet

1. Introduction

Artificial intelligence systems today are successfully applied in various fields [1]. However, the latest deep learning algorithm consists of complex configurations, and image classification systems such as CNN consist of very deep layers and large matrixs. Therefore, recently commercialized artificial intelligence algorithms are also called black box models [2]. As a result, there is an increasing demand for transparent, descriptive artificial intelligence to researchers and users of artificial intelligence systems [3]. In particular, in areas where the results of an intelligence system are critical, such as healthcare, finance, and security, the provision of compelling data on the analysis results of the system is critical [4]. In particular, in the medical field, medical diagnostic assistance systems should be informed of the interpretation of the results [5]. Therefore, it is necessary to combine the construction of a highly accurate neural network with a model that provides the analytical content.

2. Configuration of CNN Models

2.1 CNN model selection and data preprocessing

To use the CNN Internal Operation Analysis Model using Feature Map, the configuration, data preparation and learning process of the CNN-based ResNet model [6] is first required.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(No: NRF-2020R1A6A1A03040583)

For basic data for learning, use Kagle's FER2013 data-set set [7]. This data is a face image data consisting of a resolution of 48*48 and is in jpg format. A total of 28,709 data are divided into seven facial expressions and provided as learning. The total number of data used in the evaluation is 7,178.

2.2 Configuration of the ResNet Model

Prepare ResNet based on CNN form for image analysis. ResNet separates the neural network into small blocks, which are called Residual blocks. Residual blocks are characterized by the way they add the input value of the block to the output. Solve the problem of Vanishing Gradients in the neural network model deepening in this way. In this study, this approach is used and the Residual blocks are composed of overlapping. In addition, batch regularization is placed for each individual block. batch regularization normalizes the data entered into the block by means and variances, and uses it to have the same effect as 'Dropout'. The overall structure of the RESNet is shown in **Fig. 1**.

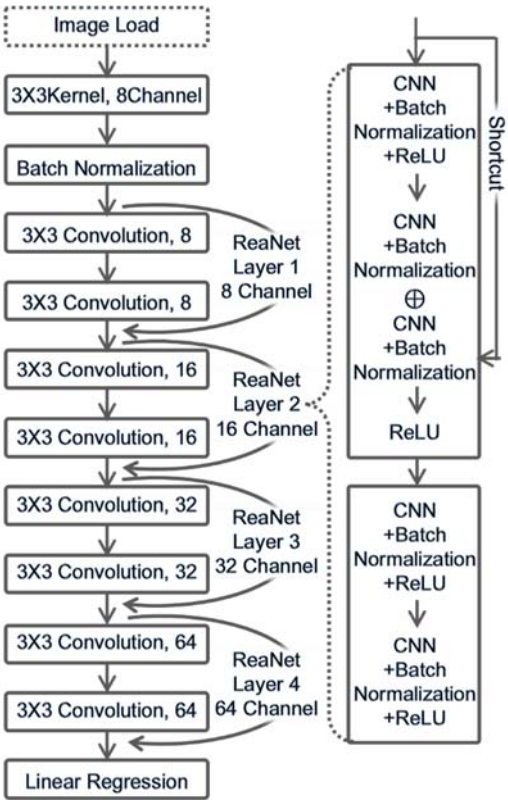


Fig. 1. RESNet structure

3. CNN Internal Operation Analysis Model using Feature Map

The form of explanation after learning and judging of neural networks uses a variety of means, including text-based, visual, and simplified explanations [8]. In particular, visual explanation is an efficient way to communicate interaction information within the model to neural network researchers and AI system users. There is LIME [9] as a visual description, and LIME displays the results of the neural network in object shapes. This approach is effective in describing object classification. However, it is less useful for understanding the internal neural network. In this study, the basic concept of LIME is approached from a different perspective. Separate images with pixels for basic implementation of the model. Invert each separate pixel. Enter the inverted pixel into the neural network. Detects changes in output nodes of the neural network. Thus, the detection results can be converted to image form and are visually expressed. Therefore, first of all, the structure of the ResNet that has been completed must be prepared. Subsequently, input the original image that requires judgement. Identify the output nodes from which the results are extracted by the final classification according to the input. Set the value MaxN for the node with the maximum value as the reference value. The next step is to compare the difference with MaxN, reversing each pixel sequentially.

Therefore, when the previous maximum node value is MaxN and the maximum value is MaxM after reversing the pixel value, the formula for determining the influence PP of pixel P, which has a positive effect on the final judgment, is as shown in Equation 1.

$$PP = Relu(MaxN - MaxM) \qquad (1)$$

ReLu(Rectified Linear Unit) has the characteristic of printing the input value as it is if the input value is less than zero and greater than zero.

Conversely, the formula for the negative impact strength of pixel NP, which has a negative effect, is as shown in Equation 2.

$$NP = Relu(MaxM - MaxN) \qquad (2)$$

In order to visually represent PP and NP values, normalization is required. Afterwards, the brightness is adjusted to 256 steps and expressed in visual images. The formula for this is like Equation 3. Pnew is the new value of the current pixel, Pmax is the maximum value of the pixel in the Feature Map, and Pmin is the minimum value..

$$Pix_{new} = \frac{Pix - Pix_{min}}{Pix_{max} - Pix_{min}} \times 256 \qquad (3)$$

The images extracted with this formula can identify the shape that has a positive or negative effect on the extraction results of the neural network. The overall operation is shown in **Fig. 2**.

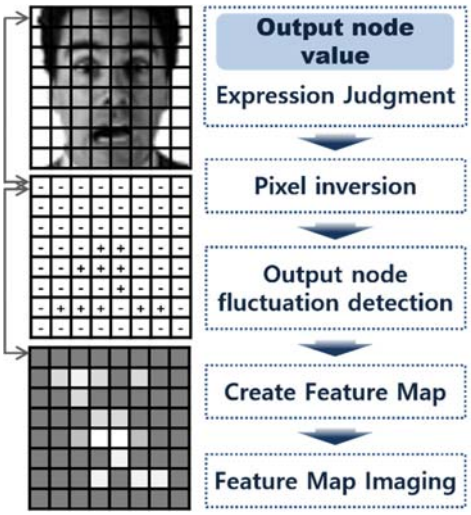


Fig. 2. Feature map imaging process

However, when performed in pixels, the characteristics of the detailed image pattern produced on layers close to the input layer are expressed in Feature Map. These images are the actual expression of the detailed operating characteristics of the neural network and help to understand the inside of the neural network. However, it interferes with understanding the operating characteristics of the hidden layer. In this case, combining the units of the input change with pixels gives a smoother result. **Fig. 3** shows the analysis results according to the amount of learning.

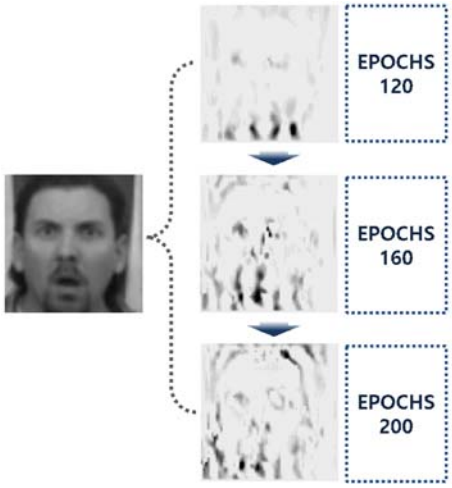


Fig. 3. Results image according to the amount of learning

As a result of image extraction, it can be seen that the working section of the neural network becomes clear as the amount of learning increases.

4. Conclusions

Currently, artificial intelligence is applied to various industries. The applied artificial intelligence algorithm is carried out in the process of big data collection, algorithm composition, and evaluation regardless of form, and finally performance is verified by evaluation results such as accuracy. Despite various evaluation figures, however, it is not known whether the actual internal direction of artificial intelligence is correct or whether it shows the right result. Thus, the combination of artificial intelligence accuracy and ancillary algorithms that can explain the interior of XAI is gradually becoming important in the research and working environment.

This paper used CNN-based ResNet for the composition of auxiliary description method, and implemented and combined XAI function using CNN Internal Operation Analysis Model using Feature Map. The implemented algorithm used a partial reverse transformation of the input data and constructed a technique to image the

resulting variation in the results. Through this, it is now possible to provide a convincing basis for the outcome of the judgment of the artificial intelligence system.

References

[1] S. J. Russell and P. Norvig, “Artificial intelligence: a modern approach. Malaysia:” *Pearson Education Limited*, 2016.

[2] D. Castelvecchi, “Can we open the black box of AI?,” *Nature*, Vol 538, No. 7623, pp. 20, 2016.

[3] A. Preece, D. Harborne, D. Braines, R. Tomsett and S. Chakraborty, “Stakeholders in Explainable AI,” *arXiv:1810.00184v1*, 2018.

[4] J. Kang, J. Baek and K. Chung, “PrefixSpan Based Pattern Mining Using Time Sliding Weight From Streaming Data,” *IEEE Access*, Vol 8, pp. 124833-124844, 2020.

[5] E. Tjoa and C. Guan, “A survey on explainable artificial intelligence (XAI): Towards medical XAI,” *arXiv:1907.07374*, 2019.

[6] K. He, X. Zhang et. al., “Deep Residual Learning for Image Recognition,” In *Proc. of CVPR*, pp. 770-778, 2016.

[7] I. J. Goodfellow, D. Erhan et. al., "Challenges in representation learning: A report on three machine learning contests," in *Proc. of Neural Information Processing*, pp. 117-124, 2013.

[8] A. B. Arrieta, N. Díaz-Rodríguez et. al., “Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI,” *Information Fusion*, Vol 58, pp. 82-115, 2020.

[9] M. T. Ribeiro, S. Singh and C. Guestrin, “"Why Should I Trust You?": Explaining the Predictions of Any Classifier,” in *Proc. of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1135-1144, 2016.

A Featured Appropriateness Bigdata-Discovery Approach for Human Resources Allocation in Business Processes

Dinh-Lam Pham[1], Hyun Ahn[1] , Kyoung-Sook Kim[1], Minjae Park[2], and Kwanghoon Pio Kim[1*]
[1]Data and Process Engineering Research Lab,
Division of Computer Science and Engineering, Kyonggi University
[e-mail: {phamdinhlam, hahn, khmjmc, kwang}@kgu.ac.kr]
[2]Division of Computer Information, Daelim University College
[e-mail: mjpark@daelim.ac.kr]
*Corresponding Author: Kwanghoon Pio Kim

Abstract

Allocating and managing human resources is one of the essential tasks of an organization. Efficient allocating takes advantage of available resources in the organization. Furthermore, it helps in reducing operating costs as well as improving the performance of employees in the organization. In this paper, we propose an approach to allocating human resources based on the appropriateness probability discovered from the workflow enactment event log. In particular, the allocation of human resources tasks in an organization will be implemented based on the appropriateness metric calculated in the discovered affiliation network and is applied to the reengineering phase of the process lifecycle. This approach will support organizations as well as businesses in planning the future business process and managing human resources more efficiently.

***Keywords*:** Human resources allocation, Process mining, Affiliation network, Workflow event log.

1. Introduction

Organizations need the appropriate resources to operate successful businesses in their manner. Each different organization has its own characteristics in terms of resource needs, however, they all share the common feature that requires human resources to operate the system.

Depending on the scope of the organization, the proportion of resources is allocated differently, but human resources are indispensable. So that allocating appropriate human resources will help saving time and production costs as well as operating costs.

In this paper, we introduce our approach to allocating human resources based on the appropriateness probability discovered from the workflow enactment event log. More precisely, we calculate the appropriateness probability of activities and performers from the discovered affiliation network to support the modeler in

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(grant number 2020R1A6A1A03040583). This work also was supported by Kyonggi University‘s Graduate Research Assistantship 2020.

order to allocate human resources in the reengineering phase of the process lifecycle.

The remaining of the paper is categorized as follows: In section 2, we discuss the concept of affiliation network. Section 3 and section 4 present our solution for calculating the appropriateness probability and allocating human resources with this metric. Finally, section 5 gives conclusions of the paper.

2. Affiliation Network

In the social network [1], there are two types of entities: actors and societies. These entities are related by affiliation with other members, and we called affiliation networks [2, 3]. To represent the affiliation network, we use the structure of the bipartite graph [4].

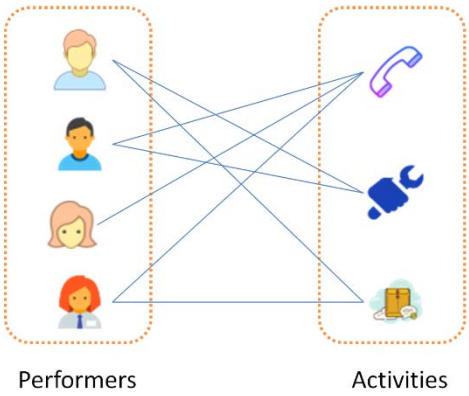


Fig. 1. Example of an affiliation network.

In the event log produced by a process-aware information system [5], we also have two types of entities: performers and activities. Activities are the actions in the system which is executed to archive a specific goal; meanwhile, performers are the person who executes the activities. It indicates that we can discover the affiliation network of performers and activities from the event log and allocating human resources based on this network.

Fig. 1 describes an example of an affiliation network. In this example, there are four performers and three activities in the network and the edges imply the activities and performers in their groups.

3. Appropriateness Probability

The discovered affiliation network brings information of operated activities and who executed these activities. Lets consider:

- Ω is a discovered affiliation network.
- n: The total activities in Ω.
- m: The total performers in Ω.
- A^iP^j: The total times activity A^i was performed by performer P^j.
- $\sum_{i=1}^{n} P^j(A^i)$: The total times performer P^j perform all activities in Ω.
- $\sum_{i=1}^{m} A^i(P^j)$: The total times activity A^i was performed by all performers in Ω.
- The appropriateness probability of activity A^i is calculated by equation:

$$AP_{A^i} = \{\varepsilon_{P^j} \mid 1 \le j \le m\}$$
$$Where: \varepsilon_{P_j} = \frac{A^iP^j}{\sum_{j=1}^{m} A^i(P^j)}$$

- The appropriateness probability of performer P^j is calculated by equation:

$$AP_{P^j} = \{\varepsilon_{A^i} \mid 1 \le i \le n\}$$
$$Where: \varepsilon_{A^i} = \frac{A^iP^j}{\sum_{i=1}^{n} P^j(A^i)}$$

By calculating the appropriateness AP_{A^i}, AP_{P^j}, we figure out the probability of performers performed an activity and the probability of activities that a performer executed. These metric values help modeler in planning and allocating human resourses in the reengineering phase of the process lifecycle in the organization.

4. Allocating and Managing Human Resources Based on Appropriateness probability

In this section, we applying our method to calculate the appropriateness probability in the Helpdesk data set [6]. First of all, we need to discover the affiliation network by using the algorithm introduced in [7]. And then, using the equations in section 3 to calculate the metrics.

Table 1 describes the appropriateness probability of performers in the activity named '*Take in charge ticket*'. To see the full result of all activities and performers, please visit this URL:

https://github.com/DinhLamPham/PropriatenessProbability

Table 1. Appropriateness probability of performers in activity '*Take in charge ticket*'

Performer	Execution Times	Appropriateness Probability
Value 2	1525	0.301383399
Value 9	603	0.119169960
Value 1	569	0.112450593
Value 13	415	0.082015810
Value 8	411	0.081225296
Value 6	306	0.060474308
Value 4	287	0.056719368
Value 16	157	0.031027668
Value 12	147	0.029051383
Value 14	146	0.028853755
Value 11	106	0.020948617
Value 10	105	0.020750988
Value 7	89	0.017588933
Value 15	71	0.014031621
Value 17	28	0.005533597
…	…	…

In **Fig. 2**, we can see the appropriateness of performers in the activity '*Take in charge ticket*'. As shown in the figure, the performers (*Value2, Value 9, Value 1, Value 13, and Value 8*) are the most appropriateness performer for executing the activity 'Take in charge ticket'.

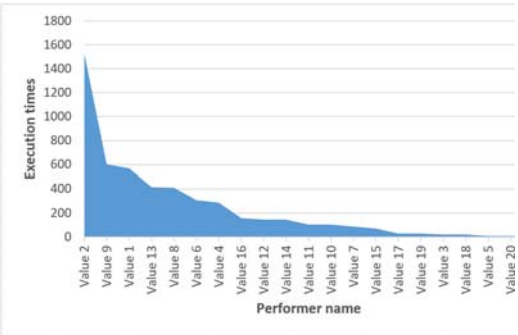


Fig. 2. The appropriateness of performers in activity named 'Take in charge ticket'

Based on this information, the system modeler can give effective decisions in allocating and managing human resources in the reengineering phase of the process lifecycle.

5. Conclusions

In this paper, we have proposed our approach to allocating human resources based on featured appropriateness probability. In this approach, the expression of appropriateness probability of activities and performers was introduced as well as applied in the Helpdesk data set. This approach promises support in planning the future business process and managing the human resources in organizations efficiently.

Acknowledgments. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (grant number 2020R1A6A1A03040583). This work also was supported by Kyonggi University's Graduate Research Assistantship 2020.

References

[1] Scott, J., 1988. Social network analysis. *Sociology*, *22*(1), pp.109-127.

[2] Lattanzi, Silvio, and D. Sivakumar. "Affiliation networks." *Proceedings of the forty-first annual ACM symposium on Theory of computing*. 2009.

[3] Luke, Douglas A. "Affiliation networks." A User's Guide to Network Analysis in R. *Springer, Cham*, 2015. 125-144.

[4] Asratian, Armen S., Tristan MJ Denley, and Roland Häggkvist. "Bipartite graphs and their applications". Vol. 131. *Cambridge university press*, 1998. 7-10.

[5] Dumas, Marlon, Wil M. Van der Aalst, and Arthur H. Ter Hofstede. "Process-aware information systems: bridging people and software through process technology". *John Wiley & Sons*, 2005.

[6] Verenich, Ilya (2016), "Helpdesk", *Mendeley Data, V1*, doi: 10.17632/39bp3vv62t.1

[7] Pham, D.L., Ahn, H., Kim, K.S. and Kim, K.P., 2019, December. "An Approach for Discovering Affiliation Network from XES Formatted Workflow Log". *In 2019 6th NAFOSTED Conference on Information and Computer Science (NICS)* (pp. 336-341). IEEE.

Traffic-Crime Association Knowledge using Text Mining in Streaming Contents

Ji-Won Baek[1], Hye-Jeong Kwon[2], Joo-Chang Kim[1], Hoill-Jung[3] and Kyungyong Chung[2*]
[1]Department of Computer Science, Kyonggi University, Suwon, South Korea
[2]Division of Computer Science and Engineering, Kyonggi University, Suwon, South Korea
[3]Department of Computer Software, Daelim University, Anyang, South Korea
[e-mail: jwbaek@kyonggi.ac.kr, rnjsgpwjd121@kyonggi.ac.kr, kjc2232@naver.com, hijung@daelim.ac.kr, dragonhci@gmail.com]
*Corresponding author: Kyungyong Chung

Abstract

Traffic crimes include drunk driving, speeding, and traffic violations. Anyone needs to understand this information easily and cope with the situation. Accordingly, in this paper, we propose knowledge related to traffic crime using text mining in streaming content. The proposed method extracts traffic accidents and traffic crime keywords through text mining from streaming content related to traffic accidents and traffic crimes. A transaction is formed for each content to discover the association rules. In the related rule, a rule that occurs frequently in traffic accidents and traffic crime keyword transactions, and satisfies the minimum support map of 0.2 and the reliability of 0.8 to extract the reliable rule is found. Accordingly, related knowledge is discovered and visualized to provide knowledge related to traffic crime.

***Keywords*:** Traffic-Crime, Text-mining, Streaming Contents, Association Knowledge

1. Introduction

In modern society, the aging population continues to increase. Accordingly, elderly drivers and elderly crimes are also increasing, and traffic crimes account for a high proportion of them. Traffic crimes are classified into traffic crimes and traffic crimes other than traffic accidents. According to the crime analysis data of the Prosecutors' Office, the age groups that account for the highest percentage of traffic offenders are between 41 and 50 years old and between 51 and 60 years old. In addition, it appears that the time zone in which traffic crimes occur most often is between 20:00 and 04:00 [1]. Various vehicles such as buses, trucks, and private vehicles exist on the road. In addition, there are differences in driving speed and driving distance depending on the driver's propensity. There are also novice drivers and experienced drivers depending on the driver's career. Since novice drivers lack experience, they are less knowledgeable about how to cope with unexpected situations and traffic accidents, and how to respond to traffic crimes than experienced drivers. There are drivers who frequently speed on the road, drivers who go slower than the minimum speed, and drivers who ignore the signal. Currently, the most common violation of road laws is the violation of the speed limit. However, there are cases where the driver thinks he or she is driving safely. Accordingly, traffic accidents frequently occur. Traffic accidents are unpredictable accidents that can occur anytime, anywhere, with minor, serious, and even fatalities. In addition, speeding on the road is difficult because the number of vehicles has increased [2,3]. Therefore, drivers who cause traffic accidents occur. In order to solve this problem, it is necessary to recognize the

This work was supported by the GRRC program of Gyeonggi province. [GRRC KGU 2020-B03, Industry Statistics and Data Mining Research]

seriousness of traffic accident crimes and enable safe driving. In this paper, we propose traffic-crime association knowledge using text mining in streaming contents.

2. Related Work

2.1 Information Extraction using Text mining

Text is the most used data type in real life. People collect and share a variety of information on news, blogs, and social media through text. However, since it is difficult to find meaningful information in a vast amount of text content, valuable information is found through text mining [4,5]. Text mining is a technology that extracts meaningful information from unstructured text data. It is widely used in document classification and information extraction. Information extraction automatically extracts information having important meaning from text content and provides it to the user. Accordingly, it is possible to briefly provide various information such as important keywords and core concepts of the text content. H. J. Kim et al. [6] proposed a graph of association knowledge using TF-IDF based Ranking Score. The proposed method generates a keyword ranking by calculating a TF-IDF (Term Frequency-Inverse Document Frequency) weight in a large amount of text content. Optimized transactions are created based on keywords that received high scores in the TF-IDF-based ranking. Accordingly, the knowledge according to the association rules is extracted using the association rules algorithm. Therefore, it provides important information in text content. In addition, document classification identifies the characteristics of each text content and classifies relevant documents with similar content or form according to the purpose [7]. Accordingly, M. J. Kim et al. [7] proposed a word embedding-based traffic document classification model to detect emerging risks using sentiment similarity weight. The proposed method is to detect new dangers by considering and classifying the importance of keywords and the polarity of keywords in traffic-related documents. This classifies traffic risk documents by expanding the scope of use of keywords by using semantically important keywords not included in the existing sentiment analysis dictionary. Therefore, it provides traffic information considering the user's location and route.

3. Traffic-Crime Association Knowledge using Text Mining in Streaming Contents

3.1 Streaming Content Collection and Preprocessing

Streaming content is being created in real-time. It is composed of various traffic accident information and traffic crime knowledge text. Streaming content collection collects traffic accidents and traffic crime-related content through crawling on KBS (Korean Broadcasting System) [8] homepages. KBS are the three major broadcasting companies in Korea and provide the most information to people. Each of the collected traffic crime-related text contents is composed of one document. This is preprocessed using text mining. The preprocessing method proceeds with the morpheme analysis to extract important keywords. Morphological analysis extracts keywords in the form of nouns by removing stop words, numbers, and punctuation marks [9]. Each document that has been pre-processed to generate related knowledge is composed of transactions. **Table 1** shows transaction of traffic accidents and traffic crime.

Table. 1 Transactions of traffic accidents and crime

TID	Traffic Keyword
TID_001	Drinking, speed, collision, black box, insurance fraud, assault, used imported cars, driving
TID_002	Accident, fog, Seohae Bridge, freight car, collision, congestion, sea, braking distance, driving
TID_003	Collision, car, freight car, traffic congestion, Seohae Bridge, accident
TID_004	Drinking, traffic accident, driver, traffic violation, getaway, Haeundae
…	…

3.2 Extraction and Visualization of Association Knowledge

There are various types of traffic crimes, such as hit and run, drunk driving, speeding, traffic violations, unsecured safety distance, and insurance fraud. This poses a danger to people's lives, bodies, and property. Each traffic crime has associated knowledge. Therefore, knowledge is

extracted from traffic accidents and traffic crime transactions using an association rule algorithm. The association rule algorithm finds hidden association rules in a large amount of data and works in large-scale transactions. It also creates hidden rules that are easy for anyone to understand. The indicators used in the association rules have support and reliability, and there is an improvement in determining the effectiveness of the rules. The support level represents the ratio of traffic Keyword_A and Traffic Keyword_B simultaneously included among all traffic accidents and traffic crime transactions. Reliability represents the percentage of transactions including Traffic Keyword_A that includes Traffic Keyword_B. In addition, the degree of improvement grasps the meaning between Traffic Keyword_A and Traffic Keyword_B. When the degree of improvement is 1, it represents an independent relationship, and when it is negative, it shows a negative correlation. When it is positive, it shows a positive correlation. **Table 2** shows the traffic crime-related knowledge bases extracted from transactions of traffic accidents and traffic crime.

Table 2 Traffic crime-related knowledge bases

Association Rules	Sup.	Conf.	lift
car, Seohae Bridge → accident	0.2	1.00	5.00
black box, speed → collision	0.2	1.00	2.50
driving, insurance fraud, insurance money → Bus	0.2	1.00	5.00
driving, insurance money, traffic accident → Bus	0.2	1.00	5.00
…	…	…	…

4. Conclusions

Traffic crime poses a danger to people's lives and property. Therefore, you need information and knowledge to cope with the risks. In this paper, we proposed knowledge related to traffic crime using text mining in streaming content. The proposed method was documented by collecting streaming data related to traffic accidents and traffic crimes through crawling. Accordingly, keywords related to traffic accidents and traffic crimes were extracted from each document through text mining. After completing the pre-processing, traffic accident and traffic crime transactions were composed. In traffic accidents and traffic crime transactions, related knowledge related to traffic crime was extracted through the algorithm of the related rules. Therefore, anyone can easily understand information on traffic accidents and traffic crimes and provide knowledge to cope with the situation. In the future, the proposed method will be upgraded to provide personalized information.

References

[1] Prosecutors Office, https://www.spo.go.kr/.
[2] D. H. Shin, K. Chung, R. C. Park, "Prediction of Traffic Congestion Based on LSTM Through Correction of Missing Temporal and Spatial Data", *IEEE Access*, Vol. 8, pp. 150784-150796, 2020.
[3] J. W. Baek, K. Chung, "Pothole Classification Model Using Edge Detection in Road Image", *Applied Sciences*, Vol. 10, No. 19, pp. 6662, 2020.
[4] J. Rashid, S. M. A. Shah, A. Irtaza, "Fuzzy topic modeling approach for text mining over short text", *Information Processing & Management*, Vol. 56, No. 6, pp. 102060, 2019.
[5] J. C. Kim, K. Chung, "Associative feature information extraction using text mining from health big data", *Wireless Personal Communications*, Vol. 105, No. 2, pp. 691-707, 2019.
[6] H. J. Kim, J. W. Baek, K. Chung, "Optimization of Associative Knowledge Graph Using TF-IDF Based Ranking Score", *Applied Sciences*, Vol. 10, No. 13, pp. 4590, 2020.
[7] M. J. Kim, J. S. Kang, K. Chung, "Word-embedding-based Traffic Document Classification Model for Detecting Emerging Risks using Sentiment Similarity Weight", *IEEE Access*, Vol. 8, pp. 183983-183994, 2020.
[8] KBS, http://www.kbs.co.kr/.
[9] A. Karami, M. Lundy, F. Webb, Y. K. Dwivedi, "Twitter and Research: a Systematic Literature Review through Text Mining", *IEEE Access*, Vol. 8, pp. 67698-67717, 2020.

Correlation Analysis between Influencing Factors and Traffic Speed in Accident Prone Areas

Min-Jeong Kim[1], Ji-Soo Kang[2], Sung-Soo Park[2], Dong-Hoon Shin[2], Roy C. Park[3] and Kyungyong Chung[1,*]
[1]Division of Computer Science and Engineering, Kyonggi University, Suwon, South Korea
[2]Department of Computer Science, Kyonggi University, Suwon, South Korea
[3]Department of Information and Communication Engineering, Sangji University, Wonju, South Korea
[e-mail: minjeog0@gmail.com, kangjs920@gmail.com, wolfboy@kyonggi.ac.kr, dhshin8227@kyonggi.ac.kr, roypark@sangji.ac.kr, dragonhci@gmail.com]
*Corresponding author: Kyungyong Chung

Abstract

As the population is concentrated in urban areas in modern society, problems such as traffic congestion and traffic accidents are becoming more severe. Accordingly, many regions of the downtown area have been selected as accident-prone areas, and the traffic safety index is also lowering. Therefore, this study analyzes the linear relationship between traffic speed increase and decrease and other traffic variables using traffic data. This extracts the main variables that affect the travel speed through regression analysis and derives a significant correlation between the extracted variables. As a result of the research, it appears that weather variables, traffic variables, and time variables affect travel speed. In this way, the statistical analysis of accident-prone areas is possible through the correlation of traffic-related variables.

***Keywords*:** Traffic speed, Regression Analysis, Correlation Analysis, Pearson correlation coefficient, Influencing factor

1. Introduction

With the development of modern civilization, 91.8% of the Korean population lives in cities, resulting in overpopulation [1]. Accordingly, the amount of traffic flowing into the urban area increases, and the number of traffic accidents is increasing. According to the Road Traffic Authority's 2019 Traffic Safety Index, the most densely populated Seoul Metropolitan City's safety index is between 70 and 80. Many ward areas are in the lowest grade with E grade [2]. Therefore, it is analyzed by sampling the many accident areas in Seoul. Many variables affect travel speed, such as the presence or absence of a traffic accident and lousy weather. This requires the selection of an independent variable that has a significant relationship with the dependent variable. In this study, the main variables affecting the speed increase and decrease of traffic in the accident-prone area are extracted through the regression analysis. It also derives and analyzes linear relationships between extracted variables. Through this, it can be used for statistical analysis for public policies implemented in accident-prone areas.

2. Related Works

2.1 Correlation analysis between variables

Computers and communication technologies are changing into an information society. This requires a method to derive meaningful

This work was supported by the GRRC program of Gyeonggi province. [GRRC KGU 2020-B04, Image/Network-based Intellectual Information Manufacturing Service Research]

relationships between numerous data. Correlation analysis is an analysis method that provides a linear relationship between two variables and is measured using a correlation coefficient. Correlation coefficients for correlation analysis are Pearson correlation coefficient [3], Spearman correlation coefficient [4], etc. The Pearson correlation coefficient is often used in continuous data to determine the relevance of two variables. This means that the values of -1 to +1 remain unchanged. The closer the correlation coefficient is to -1, the negative correlation, and the closer to +1, the positive correlation. And if the value of the correlation coefficient is 0, there is no linear correlation. Equation (1) is the Pearson correlation coefficient. $cov(x, y)$ is the covariance of x and y, and $\sigma_x \sigma_y$ is the standard deviation of x and y.

$$Pearson = \frac{cov(x,y)}{\sigma_x \sigma_y} \quad (1)$$

The Spearman correlation coefficient is used when analyzing data measured on the same scale, and the same value from -1 to +1 is used as the Pearson correlation coefficient. It means that the degree is consistent, and the correlation to -1 implies that the two variables' ranks are opposite. Equation (2) gives the Spearman correlation coefficient. n is the total number of data. x_i has the i-th data rank from x. y_i has the i-th data from y. $\bar{x}$ and $\bar{y}$ represent the mean of x, y.

$$\text{Spearman} = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^n (x_i - \bar{x})^2} \times \sqrt{\sum_i^n (y_i - \bar{y})^2}} \quad (2)$$

S.Arain et al. [5] presented the correlation analysis between drivers at traffic congestion points. This is a statistical analysis through a questionnaire, which shows that a few' behavior has a strong correlation. Through this, we derived that the behaviors between different drivers are interdependent, and based on the analysis results, recommendations for the transportation system were presented.

3. Manipulation of Virtual Model

3.1 Data Collection & Processing

The data uses traffic speed data for each road from January to September 2020, provided by the Seoul Transportation Information Center (SEOUL Topic) [5]. It consists of the date, day of the week, road name, link ID, starting point name, endpoint name, direction, distance, number of lanes, function type classification, city center/outskirts classification, area classification, number of traffic lights and hourly travel speed from 01:00 to 24:00. It also partially collects variables that can affect travel speed. Air pollution degree variables such as fine dust and ultrafine dust use Seoul Metropolitan Government atmospheric environment data [6], and meteorological information such as temperature, precipitation, wind speed, and cloud cover uses ground observation data provided by the Meteorological Agency [7]. Table 1 shows the variables that make up the collected data. In **Table 1**, str, cont and cate mean string, continuous and categorical data, respectively.

Table 1. Configuration Variables of Data

	name	data type
Traffic speed data	date	str
	day	cate
	time	cate
	road_name	str
	rink_ID	str
	Start_point	str
	end_point	str
	direction	cate
	distance(m)	cont
	road_num	cate
	road_type	cate
	center_out	cate
	district	cate
	traffic_speed	cont
	light_num	cate
Air pollution data	fine_dust	cont
	ultrafine_dust	cont
Weather data	average_temperature	cont
	min_temperature	cont
	max_ temperature	cont
	rainfall	cont
	average_wind	cont
	cloud	cont

3.2 Extraction of Major Variables using Correlation Analysis

Regression analysis [8] is conducted to find out the significant relationship between traffic speed and other variables. Regression analysis can

determine the extent to which the independent variable affects the dependent variable and is often used in statistical data analysis. The data's dependent variable is set as the travel speed, and the independent variable is composed of the remaining variables to proceed with multiple regression analysis. In regression analysis, if the p-value of coefficients is less than 0.05, it is statistically significant, and the smaller the value, the more correlated. Therefore, only independent variables whose p-value is less than 0.05 are extracted. As a result of regression analysis, distance, light_num, average_temperature, rainfall were selected as the top four variables.

The correlation analysis [9] is performed to determine the correlation between the five selected independent variables and dependent variables. Because both dependent and independent variables are continuous data, they are analyzed using Pearson correlation variables. **Table 2** shows the results of correlation analysis between major variables. In **Table 2**, represents the dependent variable, traffic speed. Y_1 The independent variables from X_1 to X_4 mean distance, light_num, average_temperature, rainfall variables.

Table 2. Configuration Variables of Data

	Y_1	X_1	X_2	X_3	X_4
Y_1	1.00				
X_1	0.74	1.00			
X_2	-0.82	0.49	1.00		
X_3	0.59	0.11	0.04	1.00	
X_4	-0.63	0.09	0.13	0.67	1.00

As a result of the correlation analysis between the dependent and independent variables, there is a positive correlation with the distance and temperature variables and a negative correlation with the number of traffic lights and precipitation. Therefore, in the accident-prone areas of Seoul, as the road's length and temperature increase, the travel speed increases, and as the number of traffic lights and precipitation increases, the travel speed tends to decrease. Through this, it can be seen that traffic-related variables and weather variables have a significant linear relationship with the travel speed

4. Conclusions

In this study, like traffic accidents and traffic congestion, significant variables that affect travel speed were extracted and analyzed for correlation. Independent variables having a significant relationship with travel speed were extracted through multiple regression. Also, the correlation between the dependent variable and the independent variable was analyzed to see what kind of correlation it had. This allows us to identify and analyze the key factors that affect the problem.

References

[1] E-NARA Index, Government Monitoring System, http://www.index.go.kr/.

[2] Korea Road Traffic Authority, https://www.koroad.or.kr/.

[3] H. Xu, and Y. Deng, “Dependent evidence combination based on Shearman coefficient and Pearson coefficient,” *IEEE Access*, Vol 6, pp. 11634-11640, 2017.

[4] P. Schober, C. Boer, and A. L. Schwarte, “Correlation coefficients: appropriate use and interpretation,” *Anesthesia & Analgesia*, Vol. 126, No. 5, pp. 1763-1768, 2018.

[5] S. Arain, F. Shaikh, and M. M. Shaikh, “Problem of Traffic Congestion and Correlation Analysis of Driving behaviors in Qasimabad, Hyderabad,” *Mehran University Research Journal of Engineering and Technology*, Vol. 36, No. 1, pp. 139-148, 2017.

[6] Transport Operation & Information Service, https://topis.seoul.go.kr/.

[7] Seoul Metropolitan Government Atmospheric Environment Information, https://cleanair.seoul.go.kr/.

[8] Korea Meteorological Administration, https://www.weather.go.kr/.

[9] Z. Sun, J. Wang, Y. Chen, et al., “Influence factors on injury severity of traffic accidents and differences in urban functional zones: The empirical analysis of Beijing,” International journal of environmental research and public health, Vol. 15, No. 12, pp. 2722, 2018.

Factors Affecting Korean Public Opinion of the Police: The Conflict Perspective

Myeonggi Hong [1], Jeong-Hyeon Chang[2], Jeeseon Hwang[3], Nayoung Ko[4], EuiGab Hwang[5*]
[1],Department of Criminology, Kyonggi University
[2]Contents Convergence Software Research Center, Kyonggi University
[3]Department of Political Science and International Relations, Seoul National University
[4],Department of Criminology, Kyonggi University
[e-mail: audrl0406@gmail.com, tiger564@nate.com , ninza4966@naver.com, asprpy@naver.com, flyhwang@gmail.com]
*Corresponding author: EuiGab Hwang

Abstract

This study aims to explain factors affecting overall confidence in the police force among Korean people based on the conflict perspective. Data for the current study was obtained from the Korea Social Integration Survey(KSIS) of 2018. 8,000 citizen survey instruments were utilized and various socio-demographic variables and conflict perspective-based variables were included in the multi-nominal logistic regression. Results revealed that conflict-based perceptions such as the implausibility of social status elevation, the unjustness of society and lack of communication with the government and congress tended to be linked to a lower level of confidence in the police. On the other hand, high levels of life satisfaction and a higher perception of the country's democracy were linked to stronger confidence in the police. Among all independent variables, perceived unjustness of society was the strongest factor in explaining low confidence in the police. Implications for policies and future research were discussed based on the findings.

***Keywords*:** Confidence in the Police, Conflict Perspective, Multinominal Logistic Regression, Instrumental and Expressive Model

1. Introduction

Confidence in the police among citizens reflects the legitimacy of the police force[1]. The fair work of the police improves the confidence of citizens and gives them legitimacy[2]. Legitimacy granted to citizens is indispensable for efficacious enforcement of the law. However, depending on the situation, it is at times inevitable for the police to take forceful action. In such a situation, legitimacy allows citizens to adapt to the coercive power of the police[3].

There exist two typical models that explain confidence in the police among citizens. The instrumental model explains that citizens' confidence in the police is linked to how adeptly the police are performing their jobs, especially regarding reducing crime and the fear of crime[4,5,8]. The expressive model forms a link between the fear of crime and confidence in the police based on a cause-and-effect relationship. The model suggests that both fear of crime and confidence in the police can be explained by public concerns about moral status and social cohesion[6,7,9,10,11].

The premise of both the instrumental model and the expressive model is that citizens regard the police as a tool for formal and informal control[12,7]. Therefore, the two models explain that the criminal justice system operates based on the consensual manner of society members joined by common understanding. On the other hand, the conflict perspective explains that the criminal justice system can be seen as an agency

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(grant number 2020R1A6A1A03040583).

working for the benefit of a ruling group[12]. Under the criminal justice system, police activities can vary from group to group. For example, citizens may think that police activities are friendly to the upper class and unfriendly to the lower class[5,14]. Such a conflict based perspective may influence confidence in the police.

An abundance of previous research explains confidence in the police upon the basis of instrumental and expressive models. However, there is a lack of previous research regarding confidence in the police which is based on the conflict perspective and is conducted within the context of South Korea To fill this void, this study conducts a Multinominal Logistic Regression analysis using data from the Korea Social Integration Survey. The results explain the effects of major variables of the conflict perspective on confidence in the police,.

2. What Influences Confidence in the Police: Conflict Perspective

In various studies of Western society, classes are divided by race and ethnicity. Especially in the United States, black people feel higher levels of discrimination in education, employment, health care, and housing. They perceive that the cause of discrimination is an unjust social structure. However, white people believe that black people have a weaker will to succeed in life[15]. This conflict based perspective also assumes that black or Hispanic people would regard the criminal justice system to be unfair[15]. Furthermore, previous research has amply shown that that minorities are discriminated against in trials and arrests[16,17].

In Asia, the theory of conflict has rarely been applied to explanations about trust in the criminal justice system based on social status or preferred political parties. Only limited research exists in this field. For example, Wu and Sun(2008) revealed that, the lower the social status of the Chinese people, the more negative they feel about criminal justice[18]. Ryu and Hong(2016) explained that Korean citizens' perceptions of the police vary depending on their political orientation. They also found that Korean citizens have different levels of confidence in the police depending on class and political tendencies, with progressive and moderate political tendencies particularly being negatively associated with confidence in the police[19].

3. Research Model

Studies related to the conflict theory differentiate between power groups and non-power groups based on the categories of class, race, gender and age. Traditionally, race has often been a factor considered from the Western context, and social fairness or political orientations have often been studied within the Asian context.

With regard to confidence in the police, the present study attempts to analyze the effects of conflict based perceptions such as perceived possibility of elevation of current status, perceived fairness of society, perceived communication with the government agencies, and political tendencies.

This study hypothesizes the following:

Hypothesis 1. *Citizen perceptions of the impossibility of class status elevation will be negatively related to confidence in the police*

.

Hypothesis 2. *Citizen perceptions of unjustness of a society will be negatively related to confidence in the police.*

Hypothesis 3. *Perceptions of lack of citizen communication with the government and the congress will be negatively related to confidence in the police.*

4. Methodology

4.1. Measures

- *Confidence in the Police*[20,21,22,6,23]
  1. How well do you believe the police are doing their job?

  The response options for each item were (1) strongly disagree, (2) disagree (3) agree, and (4) strongly agree

- *Conflict based Vvariables*
  1. Perceptions of impossibility of current class status elevation (all survey items were recoded):
     - I think I can elevate my social and economic status if I work hard in Korea.

- I think that if my child works hard in Korea, he or she can elevate his or her social and economic status

2. Perceptions of unjustness of a society: (α=0.809, all items were recoded)
   - I think our society is fair for education, employment, tax, welfare, regional balance, political activities, press, business, social distribution, military service.

   The original response option for each survey item was (1) strongly disagree, (2) disagree (3)agree (4) strongly agree, which was recoded to reflect conflict aspects.

3. Perceptions of lack of citizen communication with government and congress: (α=0.814)
   - How well do you think communication between the central government and the people is going?
   - How well do you think communication between the National Assembly and the people is going?
   - How well do you think communication between the local government and the people is going?
   - How well do you think communication between local councils and the people is going?

   The response options for each item were (1) very weak, (2) weak (3) strong (4) very strong

4. Political tendency
   - What do you think of your political tendency?

   The response options for each item were (1) Conservative (2) Moderate (3) Progressive

- *Expressive Perspective based Variables*

1. Satisfaction level of life
   - How satisfied are you overall with your life?

   The response options for each item were 0=strongly dissatisfied to 10=strongly satisfied
2. Social cohesion: (α=0.554)
   - I have a strong sense of belonging to the area where I live now.
   - How well do you think communication between neighbors is going on in our society?
   - How much do you trust your neighbor?

   The response options for each item were (1) strongly disagree, (2) disagree (3) agree (4) strongly agree
3. Democratic satisfaction: (α=0.904)
   - How satisfied are you with the current level of democracy in our country?
   - How do you think the level of democracy in our country will change in five years?

   The response options for each item were 0=strongly dissatisfied ~ 10=strongly satisfied'

4.2. Samples

The study utilized Korea Social Integration Survey (KSIS) 2018 data collected by the Korean Institute of Public Administration.

A total of *8,000* surveys were included in the final statistical analysis after cleaning missing data**.**

4.3. Analytic Models

Studies that measure the confidence in the police in a single question were mainly used for logistic regression, sequencing logistic regression, and multi-universal logistic regression[24,25,18,26,27].

5. Results

Our analyses revealed that conflict-based perceptions such impossibility of the elevation of current status, unjustness of society and lack of communication with the government and the congress tended to be linked to a lower level of confidence in the police. On the other hand, a higher level of satisfaction of living and a higher perception of the country's democracy were both linked to a higher confidence in the police. Among all independent variables, perceived unjustness of a society was the strongest factor in

explaining confidence in the police. Implications for policies and future research were discussed based on the findings[20,29].

Table 1. Multinominal Logistic Regression

	Model 1 (Conflict Model)				Model 2 (Full Model)			
	B	e^B	z		B	e^B	z	
Contrast1 **a great deal / none at all**								
Impossibility of elevation of class status	-0.327	0.721	-4.970	***	-0.260	0.771	-3.820	***
Perceived unjustness of a society	-0.311	0.733	-16.860	***	-0.286	0.752	-15.130	***
Lack of communication with governt/cong	-0.488	0.614	-11.290	***	-0.463	0.629	-10.550	***
Political tendency_progressive	-0.379	0.684	-1.670		-0.542	0.582	-2.260	*
Political tendency_moderate	0.016	1.016	0.080		-0.038	0.963	-0.180	
Satisfaction with life					0.030	1.030	0.560	
Social cohesion					0.226	1.254	3.550	***
Perceived Level of Democracy					0.113	1.119	4.060	***
Sex					-0.186	0.830	-1.110	
Age					-0.092	0.912	-1.230	
Education					0.000	1.000	0.000	
Income					-0.054	0.948	-1.300	
Contrast2 **quite a lot / none at all**								
Impossibility of elevation of class status	-0.135	0.873	-4.380	***	-0.104	0.902	-3.170	**
Perceived unjustness of a society	-0.193	0.825	-19.100	***	-0.183	0.833	-17.660	***
Lack of communication with governt/cong	-0.330	0.719	-15.170	***	-0.318	0.728	-14.380	***
Political tendency_progressive	-0.074	0.929	-0.620		-0.234	0.791	-1.830	
Political tendency_moderate	0.113	1.120	1.030		0.021	1.021	0.180	
Satisfaction with life					0.038	1.038	1.420	
Social cohesion					0.045	1.047	1.470	
Democratic satisfaction					0.054	1.055	4.010	***
Sex					-0.242	0.785	-2.740	**
age					-0.029	0.971	-0.730	
Education					0.079	1.082	1.640	
Income					-0.016	0.984	-0.720	
Contrast3 **Not very much / none at all**								
Impossibility of elevation of class status	-0.101	0.904	-3.550	***	-0.084	0.920	-2.770	**

Perceived unjustness of a society	-0.103	0.902	-11.170	***	-0.096	0.908	-10.120	***
Lack of communication with governt/cong	-0.184	0.832	-9.090	***	-0.177	0.838	-8.610	***
Political tendency_progressive	-0.058	0.944	-0.510		-0.135	0.874	-1.120	
Political tendency_moderate	0.111	1.117	1.070		0.072	1.075	0.670	
Satisfaction with life					0.035	1.035	1.410	*
Social cohesion					-0.001	0.999	-0.050	
Democratic satisfaction					0.037	1.038	3.010	**
Sex					-0.155	0.856	-1.870	
age					0.027	1.027	0.710	
Education					0.053	1.054	1.160	
Income					-0.007	0.993	-0.360	
χ^2(df)	1607.80(15)***				1678.65(36)***			
Pseudo R^2	0.096				0.100			

6. Conclusions

The findings of this study indicate that conflict-based factors such as impossibility class status elevation, perceived unjustness of society, and lack of communication with government and congress were consistently significant in explaining public opinion of the police among Korean people. On the other hand, the expressive model-based factors of social cohesion and political tendency were not consistently significant in explaining confidence in the police. Even though these findings indicate the importance of conflict based factors, previous research on confidence in the police has been focused on instrumental and expressive perspectives and not much attention has been given to conflict based approaches. Thus, it is necessary to turn our attention towards conflict based perspectives as well as expressive and instrumental models in our attempt to discover factors of public opinion of the police. Such an approach may give us helpful implications for reasonable policing by giving us a comprehensive understanding of factors of public opinion regarding policing and criminal justice activities.

References

[1] Sun, I. Y., Hu, R., Wong, D. F., He, X., & Li, J. C. (2013). One country, three populations: Trust in police among migrants, villagers, and urbanites in China. Social Science Research, 42(6), 1737-1749.

[2] Tyler, T. R. (2004). Enhancing police legitimacy. The annals of the American academy of political and social science, 593(1), 84-99.

[3] Tyler, T. R. (2006). Why people obey the law. Princeton University Press.

[4] Skogan, W. G. (2009). Concern about crime and confidence in the police: Reassurance or accountability? Police quarterly, 12(3), 301-318.

[5] Weitzer, R., & Tuch, S. A. (2002). Perceptions of racial profiling: Race, class, and personal experience. Criminology, 40(2), 435-456.

[6] Cao, L., Frank, J., & Cullen, F. T. (1996). Race, community context and confidence in the police. Am. J. Police, 15, 3.

[7] Jackson, J., & Bradford, B. (2009). Crime, policing and social order: on the expressive nature of public confidence in policing. The British journal of sociology, 60(3), 493-521.

[8] Cao, L., Frank, J., & Cullen, F. T. (1996). Race, community context and confidence in the police. Am. J. Police, 15, 3.

[9] Jackson, J., & Sunshine, J. (2007). Public confidence in policing: A neo-Durkheimian perspective. British journal of criminology, 47(2), 214-233.

[10] Jang, H., & Hwang, E. (2014). Confidence in the police among Korean people: An expressive model versus an instrumental model. International journal of law, crime and justice, 42(4), 306-323.

[11] Ren, L., Cao, L., Lovrich, N., & Gaffney, M. (2005). Linking confidence in the police with the performance of the police: Community policing can make a difference. Journal of criminal justice, 33(1), 55-66.
[12] Akers, R. L., Sellers, C. S., Jennings, W. G. (2017) Criminological Theories: Introduction, Evaluation, and Application 7th Edition, Oxford University Press.
[13] Weitzer, R., & Tuch, S. A. (2002). Perceptions of racial profiling: Race, class, and personal experience. Criminology, 40(2), 435-456.
[14] Hickman, M. J., Lawton, B. A., Piquero, A. R., & Greene, J. R. (2001). Does Race Influence Police Disciplinary Processes?. Justice research and policy, 3(1), 97-113.
[15] Hagan, J., Shedd, C., & Payne, M. R. (2005). Race, ethnicity, and youth perceptions of criminal injustice. American sociological review, 70(3), 381-407.
[16] Cochran, J. C., & Warren, P. Y. (2012). Racial, ethnic, and gender differences in perceptions of the police: The salience of officer race within the context of racial profiling. Journal of contemporary criminal justice, 28(2), 206-227.
[17] Walker, S., Spohn, C., & DeLone, M. (2012). The color of justice: Race, ethnicity, and crime in America. Cengage Learning.
[18] Wu, Y., & Sun, I. Y. (2009). Citizen trust in police: The case of China. Police quarterly, 12(2), 170-191.
[19] Ryu and Hong. (2016). Assessing a Relationship between Citizens' Political and Ideological Tendency Perception toward the Police, Crisisonomy, 12(2), 129-143.
[20] Cao, L., & Hou, C. (2001). A comparison of confidence in the police in China and in the United States. Journal of criminal justice, 29(2), 87-99.
[21] Cao, L., Lai, Y. L., & Zhao, R. (2012). Shades of blue: Confidence in the police in the world. Journal of criminal justice, 40(1), 40-49.
[22] Jang, H., Joo, H. J., & Zhao, J. S. (2010). Determinants of public confidence in police: An international perspective. Journal of criminal justice, 38(1), 57-68.
[23] Jang, H. (2014). Determinant of confidence in the police: Comparison between instrumental perspective and expressive perspective. Public Safety and Criminal Justice, 23(1), 240-261.
[24] Wu, Y., Poteyeva, M., & Sun, I. Y. (2012). Trust in police: A comparison of China and Taiwan. International journal of comparative and applied criminal justice, 36(3), 189-210.
[25] Wu, Y., Sun, I. Y., & Hu, R. (2016). Public trust in the Chinese police: The impact of ethnicity, class, and Hukou. Australian & New Zealand journal of criminology, 49(2), 179-197.
[26] Long, J. S. (1997). Regression models for categorical and limited dependent variables. Thousand Oaks, CA: Sage.
[27] Kaminski, R. J., Koons-Witt, B. A., Thompson, N. S., & Weiss, D. (2010). The impacts of the Virginia Tech and Northern Illinois University shootings on fear of crime on campus. Journal of Criminal Justice, 38(1), 88-98.
[28] Long, J. S., & Freese, J. (2006). Regression models for categorical dependent variables using Stata. Stata press.
[29] Cao, L., Stack, S., & Sun, Y. (1998). Public attitudes toward the police: A comparative study between Japan and America. Journal of criminal justice, 26(4), 279-2.

Criminal Face Recognition in CCTV video using VDSR for Crime Prevention

Nakhoon Choi, Ji-Soo Kang, Sohyun Lim, Hyunbin Kim and Heeyoul Kim
Department of Computer Science, Kyonggi University,
Gyeonggi-do, South Korea
Division of Computer Science and Engineering, Kyonggi University,
Gyeonggi-do, South Korea
[e-mail: skg4463@gmail.com, kangjs920@gmail.com, gip.shlim1017@gmail.com,
nu_start@naver.com, heeyoul.kim@kyonggi.ac.kr]
*Corresponding author: Heeyoul Kim

Abstract

Although the facial recognition field is the most important and oldest subject in computer vision research, it can still cause significant problems in the efficiency, or performance of the system, for images that contain small objects such as CCTV. In this paper, we propose a system that captures images with CCTV installed in daycare centers or kindergartens, recognizes the faces of ex-convicts stored in the database in the images, and warns them in advance. The proposed system can also be useful for analyzing CCTVs after an incident, as well as crime prevention. For this purpose, we propose a system consisting of a series of processes, commonly referred to as 'face recognition'. This process consists of pre-processing based on object detection and landmarks, standardization through embedding, and classification algorithms. In this paper, we compare the traditional computer vision algorithm with the latest deep-learning based method, and consider the most suitable method to apply to the proposed system. It will also investigate how to find the criminal's face stored in the database in real-time within CCTV video. This includes the possibility of several data amplification methods to obtain sufficient data to train deep running-based classifiers. As a result, the final goal of this study is to contribute to criminal investigation or prevention by recognizing criminals on CCTVs installed in educational institutions, including daycare centers.

***Keywords*:** face recognition, face detection, face matching, Super-resolution, VDSR, Crime prevention

1. Introduction

Since crimes against children are bad and the recidivism rate is very high, it is necessary to track criminals with similar criminal records in order to block such a problem in advance. In the case of daycare centers, CCTV installation is mandatory under Article 15-4 of the Child Care Act [1]. In recent years, various studies have been actively conducted to detect and recognize people in images using machine learning and deep learning technologies that extract meaningful information from images [2]. Therefore, it is possible to protect children from offenders by using this technology to detect that offenders are near the daycare center in the daycare CCTV video. In this paper, we propose criminal face recognition in CCTV video using VDSR for crime prevention. The purpose of this study is to prevent crime by detecting criminals in real-time in CCTV images, reporting and tracking them immediately. This method can also provide automated video analysis in criminal cases. In general, if CCTVs detect abnormal

situations such as the appearance of offenders at certain locations like daycare centers, the system recognizes the situation and reports it to the agency immediately to prevent crimes. The method of recognizing a face in an image consists of several processes, and a method using a traditional computer vision algorithm or a method based on deep learning can be used in each process [3,4]. In facial recognition, existing studies tend to rely on traditional computer vision algorithms, therefore they tend to be less robust to realistic image characteristics such as race or occlusion, etc. In addition, due to the characteristics of CCTV images, problems with small objects are encountered. Therefore, in this study, while considering the characteristics of CCTV images, we intend to design a model suitable for the face of Asians. Through this, we propose a system that better detects criminals.

2. Related Works

2.1 General face recognition

Face recognition is a series of processes that detect and match faces [4]. It consists of face detection, which detects a face to be recognized from image data, and face matching, which compares the detected face with the database's face. For detection, there are methods using traditional computer vision algorithms such as Haar and HOG, and methods using deep learning networks [2-4]. Methods using existing traditional computer vision techniques have the advantage of high computational speed, but have a limit of lower accuracy compared to deep learning techniques [3]. Methods using deep learning networks are being actively researched, and large-scale data and challenges have been established. Face detection based on deep learning is similar in principle to general object detection, and has an advantage of improved performance. However, it has a disadvantage that requires a lot of computation and computing power [4]. Face matching determines whether there is a person identical to the face detected in the face database. If the same face as the detected face exists in the database, the corresponding label is output to judge who it is. The operation principle is to extract facial features and landmarks from detected faces and find data with similar features and landmarks in the face database. In this process, a classifier such as KNN (K-nearest neighbor) or SVM (Support Vector Machine) is used.

2.2 Challenges for criminal face recognition in CCTV

There are several challenges in the task of recognizing criminals' faces in CCTV images. The first challenge is the small object problem. Due to the characteristics of CCTV images, the detected face object is very small and of low-resolution. This makes it difficult to extract landmarks from the detected face. **Fig. 1** shows the example of the small object problem in CCTV video. The second challenge is face occlusion where part of the face is covered by a hat, mask, or sunglasses. Even if part of the face is covered, face detection is possible. However, there is a limit that face matching becomes difficult because the features of the face cannot be properly extracted. The third challenge is the race-biased problem of open data. Existing open data is biased toward Westerners' faces, resulting in a data imbalance problem with a small number of Asian faces. Since the target criminal is Asian, if the train data is biased on the Westerner's face, a bias occurs in the train result [5]. Therefore, it is necessary to organize the data to be suitable for Asian face detection.

Fig. 1. Example of the small object problem in CCTV video

3. Detecting criminal faces in CCTV video using VDSR for Crime Prevention

Since the purpose of this study is face recognition in CCTV images, facial data is required for artificial neural network learning. The artificial neural network must be able to learn to

accurately extract facial feature points from facial data, and apply to CCTV images. In addition, it should be possible to infer that the facial features of the person extracted from the CCTV image are a specific person existing in the sex offender database. **Fig. 2** shows the flow of criminal face recognition in CCTV video using VDSR for crime prevention.

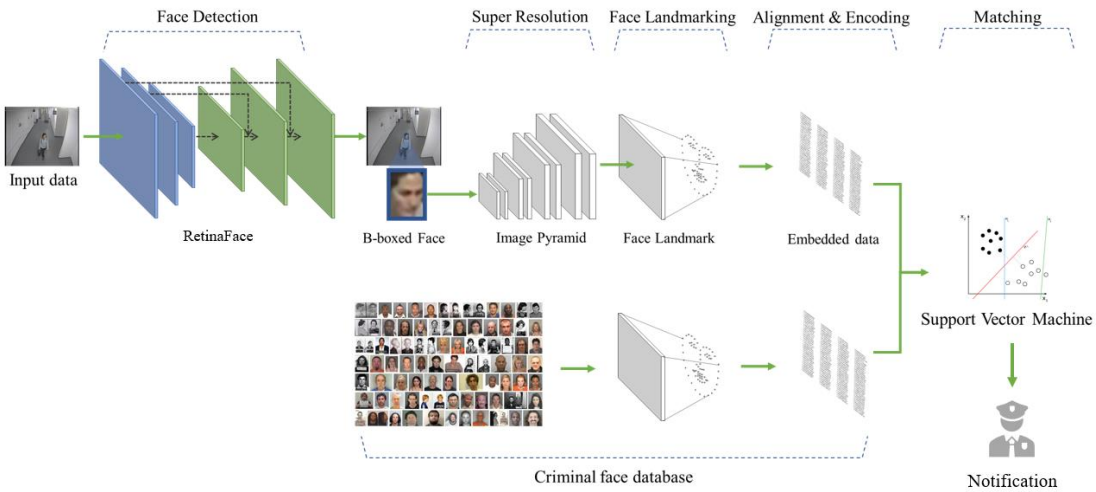


Fig. 2. Flow of criminal face recognition in CCTV video using VDSR for crime prevention

3.1 Data collection and preprocessing

The existing open data for face recognition is focused on the faces of Westerners [5]. Because of the distribution of data, when training a model, there are cases where it is difficult to detect Asian faces well. Therefore, in this study, in order to detect the faces of Asian criminals well, KFACE Dataset, a Korean face database suitable for Asian characteristics, is used [6]. The KFACE Dataset is data from a total of 1000 people and includes about 30,000 images per person reflecting 20 types of angles, 30 types of illumination, 6 types of accessories (cover), 3 types of facial expressions, and 3 types of resolution. Since the actual sex offender data we should be detected is composed of 3 pictures of left, right and front, KFACE also uses left, right and front pictures. Data augmentation is performed using left, right, and front photos to extract sufficient training data.

3.2 Face detection and super-resolution for restoring low-resolution face image

It is essential to extract feature points and landmarks that enable face recognition in CCTV images. At this time, due to the characteristics of CCTV images, a small object detection problem occurs where the size of the face object is very small. Even when detected, since the face is very small, the shape of the face is blurred, making it difficult to extract landmarks. Therefore, to solve this, super-resolution technique is used. This helps to facilitate the extraction of landmarks by restoring low-resolution face photos in high resolution.

Face detection uses the RetinaFace network [7]. RetinaFace is a network that learns face features and improves face localization performance through multi-task learning. Multi-task learning consists of face classification, face box regression, facial landmark regression loss, and dense regression loss, and through each learning, it is possible to solve the small object detection problem in CCTV images. The face detected through this process is a low-resolution image. Since it is difficult to extract facial features from low-resolution images, we use super-resolution, a technique that creates high-resolution images from low-resolution images. VDSR (Very-Deep Super-Resolution) is a representative method of super-resolution [8]. VDSR is a network that generates a residual image from a low-resolution image by learning the mapping and residual image between a low-resolution image and a high-resolution image, and generates a high-resolution image based on this. The VDSR

obtains the brightness channel of the face image detected in the CCTV image and generates a residual image for the low-resolution image. By using the generated residual image to generate a high-resolution image, a face detected as a small and low-resolution CCTV image is restored to a high-resolution image. Therefore, it is possible to improve the performance of landmark extraction, which is difficult to extract from the detected low-resolution face.

3.3 Face matching for crime prevention

The face detected in the CCTV image is super-resolution processed and then the facial landmark is extracted from this. Since the face detected in the CCTV image is not always front, it is corrected in order to accurately recognize the criminal regardless of the angle of the face. Therefore, Posing and Projection are performed in the extracted facial landmark. It arranges through Rotate, Rescale, Shear in order to locate in the image center. Embedded data is created for effective data matching of the image in which preprocessing is completed through the encoding of the facial landmark. The image-image comparison operation of face is replaced with embedded data which becomes through encoding. This converts complex raw data of a photograph or image into a numeric encoding format, thereby reducing the amount of calculation in the comparison process and improving accuracy. We perform the above process through a pre-trained deep convolutional neural network [9], and the criminal face database is also stored as embedded data.

In order to estimate that the feature of the face extracted from the image of CCTV exists in the criminal database, face matching is performed. Comparison between embedded data encoding the feature of the face extracted from the image of CCTV and the embedding criminal database which previously encodes are performed. We perform data matching using the SVM(Support Vector Machine) classifier. The SVM classifier is a linear detector, and outputs one of the data learned about the embedded data comparison operation. If there is embedded data that matches the criminal database, output the label and notify the administrator.This allows us to find criminals in CCTV footage and prevent crime. **Fig. 3** represents the result of face matching.

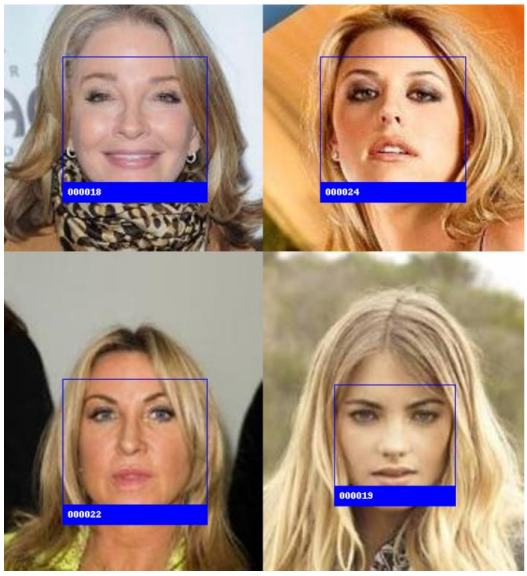


Fig. 3. Result of face matching

4. Conclusions

In this paper, we proposed criminal face recognition in CCTV video using VDSR for crime prevention. This aims to prevent crime by recognizing the face of the criminal in real time in CCTV images. To improve the performance of face detection for face recognition, we use RetinaFace, a SOTA technology, and VDSR to solve the low-definition problem caused by small object problems. In addition, the proposed model designs the model suitable for Asians using Asian face data, thereby solving the problem that the existing models were biased to Westerners and the performance of Asian detection was low. Therefore, this study helps criminal investigation and contributes to crime prevention by recognizing criminals in CCTV of daycare centers.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(grant number 2020R1A6A1A03040583).

References

[1] Child care act, Available online: http://www.law.go.kr (accessed on 18 Oct. 2020).

[2] Zhang, Suzhi, Yuhong Wu, and Jun Chang. "Survey of Image Recognition Algorithms." *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*. Vol 1, pp. 542-548, 2020. https://doi.org/10.1109/itnec48623.2020.9084972 (accessed on 23 Sep. 2020).

[3] Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*. Vol 1, pp. 886-893, 2005.

[4] Masi, Iacopo, et al. "Deep face recognition: A survey." *2018 31st SIBGRAPI conference on graphics, patterns and images (SIBGRAPI)*. IEEE, pp. 471-478, 2018. https://doi.org/10.1109/sibgrapi.2018.00067 (accessed on 23 Sep. 2020).

[5] Kärkkäinen, Kimmo, and Jungseock Joo. "Fairface: Face attribute dataset for balanced race, gender, and age." *arXiv preprint arXiv:1908.04913* (2019).

[6] KFACE Dataset, Available online: https://github.com/k-face/k-face_2019 (accessed on 23 Sep. 2020).

[7] Deng, Jiankang, et al. "RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, pp. 5203-5212, 2020.

[8] Kim, Jiwon, Jung Kwon Lee, and Kyoung Mu Lee. "Accurate image super-resolution using very deep convolutional networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 1646-1654, 2020.

[9] OpenFace Lua Script, Available online: https://github.com/cmusatyalab/openface/tree/master/models/openface (accessed on 23 Sep. 2020).

Can Explanations for Online and Offline Delinquency be Mutually Exclusive? : A Multinominal Logistic Regression of Self-Control and Opportunity Factors

Nayoung Ko[1], Myeonggi Hong[2], Jeeseon Hwang[3], Jeong-Hyeon Chang[4], EuiGab Hwang[5*]
[1],Department of Criminology, Kyonggi University
[2],Department of Criminology, Kyonggi University
[3]Department of Political Science and International Relations, Seoul National University
[4]Contents Convergence Software Research Center, Kyonggi University
[5],Department of Police Administration, Kyonggi University
[ninza4966@naver.com,audrl0406@gmail.com,tiger564@nate.com,flyhwang@gmail.com]
*Corresponding author: EuiGab Hwang

Abstract

This study examines the similarities and differences between the causes of juvenile delinquency in online and offline environments and in personal characteristics. The study utilizes data from the '2014 Survey on Juvenile Victimization in Korea.' The population of this survey is students attending middle and high schools across the country. While this paper is based on the Self-Control theory, opportunity factors based on the Routine Activity theory and the Situational Action theory are also applied. Results show that the causes of offline delinquency are low self-control, routine activity and frequent gaming and SNS use. The causes of online delinquency are high self-control, existence of communication with unknown persons and the spectrum of personal information online. The common elements of offline and online delinquency are delinquent peers. These results show that while online and offline delinquency cannot be explained with the same methodology, at the same time they are not mutually exclusive.

Keywords*:* Online delinquency, Offline delinquency, Self-Control, Opportunity factors, Routine Activity theory, Situational Action Theory

1. Introduction

The Internet has developed to the extent that it breaks down the boundaries between reality and virtual space. Accordingly, traditional crime and delinquency have begun to proliferate within the boundaries of cyberspace and the Internet. One explanatory factor is that crimes that used to take place offline have been moved online. Moreover, current issues such as the "Nth Room" incident and hidden camera crimes show that cyber crime becomes an extension of offline crime. This reflects how blurred the boundary between cyberspace and real space is.

Cyber crime is particularly dangerous. One perpetrator can attack an unspecified number of individuals, expanding the magnitude of victimization. Crime committed from remote locations is also more difficult to respond to, and anyone with internet access can commit these crimes. Cyber crimes using sophisticated technologies such as hacking are also becoming threats.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(grant number 2020R1A6A1A03040583).

Adolescents are particularly at risk. Researchers estimate that approximately 95% of adolescents use the Internet every day[8]. Dependence on electronic gadgets and the internet is thus becoming a serious problem[10]. For many adolescents, cyberspace has become an indispensable part of their everyday lives. Thus establishing solid online ethical standards for juveniles is a particularly urgent issue. However, when the causes of online and offline delinquency are not clear, it is difficult to decide which foundations to base judgements about online ethics on. This study aims to establish whether factors affecting offline and online delinquency can be mutually exclusive and to use this as a foundation for further studies on the establishment of online ethical standards for juveniles.

2. Theoretical Background

2.1 Self-Control Theory

"The Self-Control theory (1990)" posits that the difference in levels of self-control is an explanatory factor for why people do not commit crime[3]. Within this theory, 'self-control' refers to the ability to resist criminal acts. It is formed in childhood, and once established, it does not change throughout one's whole life [3, 6]. In this theory, 'low self-control' signifies a lasting personal tendency that ignores long-term costs and desires immediate, easy, and short-term rewards[3]. A major characteristic of people with low self-control is that they often chase immediate gratification and are impulsive [3]. Therefore, those with low self-control have a higher propensity for committing a crime. However, low self-control does not always lead to a person committing a crime. Low self-control can be counteracted with certain circumstances and therefore does not always lead to crime.

2.2 Self-Control and Opportunity

Gottfredson and Hirschi(1990) did not directly mention of the impact of opportunity on the relationship between self-control and crime, yet they distinguished between crime and criminality and concluded that the actual occurrence of crime is shaped by a number of "necessary conditions", including "activity, opportunity, adversaries, victims, [and] goods" [3, 7].

Current studies have analyzed the effects of opportunity factors on self-control [4]. Therefore, this study applies the Routine Activity theory [2] and Situational Action Theory[11] to conceptualize and measure opportunity factors.

The Routine Activity Theory combines offender and victim with temporal and spatial environment[1] and applies a situational approach to victimization[5]. Usually, the Routine Activity theory is used to study victimization. However, Osgood, Wilson, Malley, Bachman and Johnston (1996) extended the Routine Activity theory perspective to offenders. The routine activities or situations conducive to deviance are (1) time with peers (2) the absence of authority figures (guardian) (3) unstructured activities[9]. Opportunity factors are unstructured activities with only peers and with no guardian presence.

The Situational Action Theory (SAT)[11] aims to explain why crime happens, and more broadly why people follow and break common rules of conduct[12, 14, 15]. According to the SAT model, action is a result of situation. Namely, when person (crime propensity) and some environmental settings interact each other, action arises as a consequence [13].

This paper adopts the concept of opportunity factors introduced by Hong (2019), which included unstructured activity without parent surveillance, exposure to delinquent peers and low collective-efficacy[20]. This paper uses unstructured activity without parent surveillance and exposure to delinquent peers as opportunity factors.

2.3 Characteristic of online and offline delinquency in Korea (ROK)

According to the Judicial Yearbook (2020) [19], the number of juvenile probation cases has gradually declined since 2012, with 33.301 cases in 2018. However, the number of cases increased sharply to 36,576 in 2019 [17, 18, 19]. Juvenile delinquency is becoming an extremely urgent issue.

The report from the KIC (Korea institution of Criminology) in 2014 on juvenile cyber bullying shows that the percentage of teenagers who had experiences with both victimization and assault in cyberbullying was high, at 13.1%. Furthermore, school violence was closely related to cyberbullying. School violence in offline spaces leads to violence in cyberspace[16].

However, prior studies on juvenile delinquency tend to separate the concepts online and offline delinquency as subjects of analysis.

Using terms such as 'cyber self', these studies tend to assume disparate identities in the real world as opposed to the virtual world or to define these two types of delinquency as independent from each other.

3. Methodology

We used data from the '2014 Survey on Juvenile Victimization in Korea'. The population of this data is comprised of 7,109 middle school (grade8) and high school (grades 9,10 and 11) students across the country attending middle and high schools across the country, and the sample demographic was male (47.1%, 3347) and female (52.0%, 3762). Their age range is from 13 to 18, and the numbers of each age are 13 (1.7%, 124), 14 (25.6%, 1820), 15 (24.0%, 1705), 16 (24.8%, 1763), 17 (23.2, 1650), 18 (0.7%, 47). The data gathered was analyzed using multinominal logistic regression in SPSS.

The dependent variables are four types of delinquency: non-delinquency (4955, 69%), only offline delinquency (1058, 14.9%), only online delinquency (670, 9.4%) and commit online and offline delinquency (both) (426, 6.0%).

We examined independent variables concerning self-control factors (Low self-control), routine activity factors and situational action theory (Routine activity, delinquent peers, communication with unknown persons, exposure of personal information, gaming and SNS use).

4. Results

Table. 1. Multinominal logistic regression self-control and opportunity factors with four types of online and offline delinquency

	Only offline delinquency(A)				*Only online delinquency(B)*				*Both(C)*			
	B	S.E.	Exp(B)		B	S.E.	Exp(B)		B	S.E.	Exp(B)	
a	0.215	0.080	1.240	*	-0.262	0.096	0.770	*	0.295	0.120	1.344	*
b	0.659	0.059	1.934	***	0.113	0.070	1.120		0.671	0.087	1.955	***
c	1.399	0.071	4.052	***	0.475	0.099	1.608	***	1.487	0.097	4.425	***
d	0.122	0.080	1.130		1.115	0.088	3.048	***	0.959	0.115	2.610	***
e	0.013	0.011	1.013		0.078	0.013	1.081	***	0.105	0.017	1.111	***
f	0.044	0.021	1.045	*	0.023	0.025	1.024		0.067	0.030	1.070	***
g	0.116	0.019	1.123	***	-0.054	0.022	0.947	*	-0.005	0.028	0.995	*
h	0.644	0.092	1.903	***	-0.553	0.106	0.575	***	0.181	0.134	1.199	
i	0.114	0.033	1.121	**	0.040	0.037	1.041		-0.029	0.048	0.971	
j	0.011	0.019	1.011		-0.005	0.021	0.995		-0.056	0.027	0.946	*
k	0.442	0.214	1.556	*	0.619	0.213	1.857	*	1.348	0.208	3.849	***
l	0.438	0.126	1.549	**	0.660	0.133	1.935	***	1.006	0.146	2.736	***

a: Low Self-Control b: Routine activity c: Delinquent peers d: Communication with the unknown e: Exposure of personal information f: gaming g: SNS use h: sex i: age j: income k: Online victimization l: Offline victimization

Cox and Snell R^2	*0.251*
Nagelkerke R^2	0.297
-2LL	11112.111
x^2	2052.025***

	Only online delinquency(D)			
	B	S.E.	Exp(B)	
a	-0.477	0.115	0.621	***
b	-0.546	0.084	0.579	***
c	-0.924	0.106	0.397	***
d	0.992	0.109	2.697	***
e	0.065	0.016	1.067	***
f	-0.020	0.030	0.980	
g	-0.170	0.027	0.843	***
h	-1.197	0.129	0.302	***
i	-0.074	0.046	0.928	
j	-0.015	0.026	0.985	
k	0.177	0.251	1.193	
l	0.223	0.159	1.249	

Table. 2. Descriptive statistics

N=7109	Min	Max	Mean	Std.Dev.
Group	1.00	4.00	1.5171	.89323
Low self-control	1.00	4.00	2.4004	.48937
Routine activity	1.00	4.00	2.1031	.69228
Delinquent peers	.00	5.00	.4808	1.02547
Communication with the unknown	.00	1.00	.3490	.47669
Exposure of personal information	.00	14.00	5.2701	3.49164
Gaming	1.00	8.00	3.4413	2.07196
SNS use	1.00	8.00	3.4515	1.76903

There are clearly some differences between the causes of offline and online delinquency. First, low self-control, routine activities and delinquent friends have effects on the offline delinquent group and both online/offline group. However, they do not affect exclusively online delinquency. Moreover, the results show that the higher the self-control, the higher the possibility of online delinquency. These findings contradict previous discussions on the relationship between self-control and delinquency.

Second, communication with unknown persons online and exposure of personal information have an effect on both theonly online delinquent group and both online/offline groups.

However, SNS and games which are only available online both affect offline delinquency. Moreover, delinquent peers which can be found offline play a part in causing online delinquency. Therefore, online and offline delinquency are not mutually exclusive and must not be seen as completely different types of delinquency.

Our results show that only offline and only online delinquency both exist. Thus some adolescents act completely differently in online and offline environments, suggesting the possibility of a "cyber self". However, this term implies that one person can have two identities. The "self" is a continuous concept, and for adolescents whose daily lives are connected with online spaces it is meaningless to separate the two "selves." Therefore, it is more important to study the effects of cyberspace on personal characteristics rather than focus on the process of the formation of a "cyber self". Online ethical standards for adolescents have not yet been established. However, the idea that identity is formed not just in one realm but within the interaction of online and offline environments provides an important basis for future discussion. Standards for online ethics must not be formulated separately from ethics that affect identity in the offline world. The interplay of various online and offline factors in forming a comprehensive ethical standard is a subject which requires further discussion and study.

References

[1] Britt, C. L.. "Self-Control, Group Solidarity, and Crime: An Integrated Control Theory". *Control Theories of Crime and Delinquency*. 161-178. 2003.

[2] Cohen, L. E., & Felson, M. "Social change and crime rate trends: A routine activity approach". *American sociological review*. 588-608. 1979.

[3] Gottfredson, M. R., & Hirschi, T.. "A general theory of crime". *Stanford University Press*. 1990.

[4] Hay, C., & Forrest, W.. "Self- control theory and the concept of opportunity: The case for a more systematic union". *Criminology*, 46(4), 1039-1072. 2008.

[5] Hindelang, Michael J., Michael R. Gottfredson, and James Garofalo. "Victims of Personal Crime: An Empirical Foundation for a Theory of Personal Victimization". *Cambridge, MA: Ballinger*. 1978

[6] Hirschi, T.. "Self-control and crime". *Handbook of self-regulation*, 537-552. 2004.

[7] LaGrange, T. C., & Silverman, R. A.. "Low self- control and opportunity: Testing the general theory of crime as an explanation for gender differences in delinquency". *Criminology*, 37(1), 41-72. 1999.

[8] Lenhart, A., Madden, M., Smith, A., Purcell, K., Zickuhr, K., & Rainie, L.. Teens, "Kindness and Cruelty on Social Network Sites: How American Teens Navigate the New World of" Digital Citizenship"". *Pew Internet & American Life Project*. 2011.

[9] Osgood, D. W., Wilson, J. K., O'malley, P. M., Bachman, J. G., & Johnston, L.D.. "Routine activities and individual deviant behavior". *American sociological review*, 635-655. 1996.

[10] Semenza, Daniel C.. "Delinquency and the Digital Domain: The Influence of Online Activities and Victimization Experiences on Offline Crime and Cyber Aggression

among Adolescents". *Emory University*. ,2013

[11] Wikström, P. O. H.. "Crime as alternative: Towards a cross-level situational action theory of crime causation". *Beyond empiricism: Institutions and intentions in the study of crime*, 13, 1-37. 2004.

[12] Wikström, P. O. H.. "Individuals, settings, and acts of crime: Situational mechanisms and the explanation of crime". *The explanation of crime: Context, mechanisms and development*, 61-107. 2006.

[13] Wikstrom, P. O. H.. "Crime propensity, criminogenic exposure and crime involvement in early to mid adolescence". *Monatsschrift fur Kriminologie und Strafrechtsreform*, 92(2-3), 253-266. 2009.

[14] Wikström, P. O. H.. "Explaining crime as moral actions". *In Handbook of the Sociology of Morality (pp. 211-239). Springer, New York, NY*. 2010.

[15] Wikström, P. O. H.. "Why crime happens: A situational action theory". *Analytical sociology: Actions and networks*, 74-94. 2014.

[16] Lee, C., Sin, N., & Ha, E.. "Korea Juvenile Research Institute Press". 2014

[17] Ministry of Justice of Korea. "Judicial Yearbook". 2012.

[18] Ministry of Justice of Korea. "Judicial Yearbook". 2018.

[19] Ministry of Justice of Korea. "Judicial Yearbook". 2020.

[20] Hong, M., Hong, S., & Lee, M.. "Self-control and Opportunity as Factors of Juvenile Delinquency: Integration of Individual and Environmental Opportunity Factors". *Korean Criminological Review*, 30(4), 203-240.,2019

An Evaluation of a Deep Learning-based Vehicle Tracking Method Using Live Traffic CCTV Data

Sunwon Kang and Byoung-Dai Lee*
School of Computer Science and Engineering, Kyonggi University
Suwon, Kyonggi-do, South Korea
[e-mail: k20201101113@kgu.ac.kr, blee@kgu.ac.kr]
*Corresponding author: Byoung-Dai Lee

Abstract

Recently, due to the drastic increase in number of vehicles, the demand for automatic vehicle tracking systems with acceptable performance is increasing more than ever. Various deep learning-based approaches have been proposed and showed superior performance improvement compared to existing image processing-based solutions. In this paper, we evaluate an automatic vehicle tracking solution that utilizes latest object detection and tracking methods using live traffic CCTV data and derive directions for future improvement.

***Keywords*:** Automatic Vehicle Tracking System, CCTV, Deep Learning, Object Detection

1. Introduction

In recent times, local governments have installed nearly 510,000 CCTV surveillance cameras to prevent crimes and build safety nets. Therefore, vehicle tracking using CCTV cameras to detect escape routes for criminals is becoming increasingly important. In one case, a teenage car thief who drove a vehicle without a driver's license in Hwasun-gun, Jeollanam-do, was arrested on September 17th 2020 on the basis of vehicle tracking using CCTV cameras. Furthermore, vehicle tracking that uses highway CCTV systems has become essential for promptly apprehending numerous criminals who are vulnerable to escape, such as hit-and-run drivers and kidnappers. To address this, various image processing methods such as a saliency-based method, a method utilizing edge information, and a difference-of-frames method were proposed. However, these methods have limitations that are hard to overcome owing to which it is difficult to produce reliable results using these image processing methods. Therefore, in this study, the deep learning technology, which has recently been in the spotlight, was applied to the vehicle detection and tracking project, and it was found to outperform the image processing methods and provide desired results.

2. Related Work

Existing vehicle detection and tracking models can be classified as either models using the image processing methods or those utilizing the machine learning-based methods. The image processing method is described as follows: First, the saliency-based method[1] uses the features of a given CCTV image that stands out visually to locate the regions of the image, which are likely to contain vehicles. Next, vehicles are sorted out from the image based on the size of these regions.

This research was supported by Basic Science Researh Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (grant number: 2020R1A6A1A03040583)

Furthermore, there is a method that utilizes the edge information of an image[2]. This method detects vehicles in a given CCTV image using the boundaries between different colors within the image. The third method uses the difference-of-frames technique[3]. This method analyzes many frames of a CCTV video to detect moving objects in the image. Subsequently, these objects are classified based on their size. Lastly, the background subtraction method uses the background image without any vehicles as a reference image to detect moving vehicles in the CCTV images with the same background [4]. Although these methods have made significant contributions to the vehicle detection and tracking technology for identifying the escape routes of criminals, they have produced less than satisfactory results for the following reasons: First, these methods are sensitive to noise. Second, they could not detect vehicles whose color was similar to that of the roads. Third, the recognition rate of objects that have not moved for a certain period was low. The machine learning-based methods are as follows: First, a method that uses the histogram of oriented gradients (HoG) can recognize various types of objects, such as vehicles, objects, and people[5]. However, if support vector machines are used as the classifier, this method can only distinguish a single type of objects as true or false. Second, there is a method that utilizes the features of local binary patterns (LBP) [6]. This method performs object detection by learning the features of various patterns in the preset samples. However, it has a limitation of only using simple black and white patterns.

3. Method

In this study, the DeepSORT model[7] was used to detect vehicles in CCTV images and classify them into three classes/categories: car, truck, and bus. A unique ID was assigned to the detected vehicle, and the vehicle that is deemed to be identical was tracked frame by frame. DeepSORT model is based on the simple online and real-time tracking (SORT) model[8]. The SORT model is a tracking model that has combined detectors, the Kalman filter, and the Hungarian algorithm. The key concept of the Kalman filter is to the use available detections and previous predictions to successfully estimate the current state while maintaining the possibility of errors in the process. The relevant values are used to estimate the probability in the following order, where the probability cannot be calculated. First, the current value is predicted using the past values. Second, the actual value is predicted after applying the noise to both the predicted value and the measured value. Third, the predicted actual value is used for the next measurement. The Kalman filter was selected as the main technique for the SORT model because it is most suitable for a linear system with the Gaussian process. However, the Hungarian algorithm is an algorithm designed to solve the problem of optimization for the allocation of minimum cost. The Hungarian algorithm is used by the SORT model to perform data association.

Lastly, the DeepSORT is a model that has applied deep learning features to the SORT model. The deep learning feature was incorporated to overcome the limitations of the Kalman filter and helps to detect overlapping objects and track them individually. The overall structure of the DeepSORT model is shown in **Fig. 1**.

As shown in Fig. 1, YOLOv4[9] was selected as the object detection network for the DeepSORT model in this study. You only look once (YOLO), released prior to Yolov4, had persistent issues with the detection of small objects. To overcome this problem, a larger resolution of input was used in Yolov4. Input resolution of 224 or 256 was previously used to train the network; however, input resolution of 512 was used for training in Yolov4.

In addition, a CSPNet-based backbone was designed and used for faster machine learning. The cross stage partial network (CSPNet) structure was proposed to minimize the loss of accuracy [10]. In addition, the feature map was divided into two parts, and one part was excluded in the computation, that was combined at the back end. Using this technique, both the cost of inference and the cost of memory were reduced. Lastly, the loss of accuracy was minimized by splitting the gradient flow. Yolov4 has selected CSPNet-based CSPDarkNet53 as its backbone, and the overall structure of Yolov4 is shown in **Fig. 2**.

4. Experiments and Results

Both the DeepSORT model and Yolov4 were implemented using Windows 10-based Python

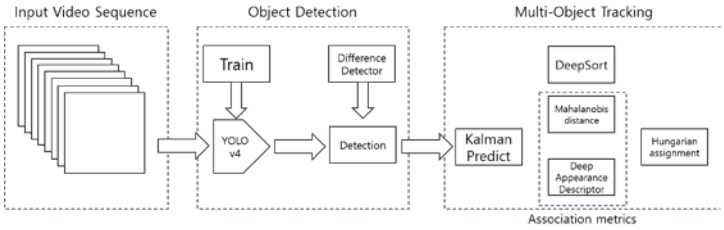


Fig. 1. The overall structure of the DeepSORT model

3.6 and TensorFlow GPU version.
The GPU used in the experiment was NVIDIA GeForce RTX 2080ti, CUDA version 10.2, and CuDNN version 7.6.5 were used. For Yolov4, the batch size was 32, and the experiment was conducted using the input image size of 608x608. In this experiment, 108000 iterations were run. In addition, the value of the weight was obtained in the units of 1000, and the test was performed using the weight with the smallest loss value. For the DeepSORT model, the batch size was 1, and the experiment was conducted using the input image size of 416x416.

4.1 Dataset

In this study, the DeepSORT model was utilized to perform object tracking in CCTV videos from three different roads (Sanmakgongdan entrance 3-way intersection, Seoksan bridge 4-way intersection, Deokgye 4-way intersection). Before performing object tracking, the Yolov4 network was trained to detect objects.
The CCTV video footage from Sanmakgongdan entrance 3-way intersection was clean and clear, and it did not contain any congested segment. The CCTV video footage from Seoksan bridge 4-way intersection was also clean and clear, but it contained a congested segment. Lastly, the CCTV video footage from Deokgye 4-way intersection was clean and clear, and it contained relatively large objects. These videos are shown in Fig. 3, and each video is labeled 'trainval video 1', 'trainval video 2', and 'trainval video 3' in the order. Furthermore, the number of training data and validation data used for each video is shown in **Table 1**. All the data consist of three classes: car, truck, and bus, and they are used after converting them from the existing PASCAL dataset format to the COCO dataset format.
Four test videos were used for object detection. The first video does not contain any congested segment, and it demonstrates a good picture quality. The second video has a reasonable picture quality and does not contain any congested segment, but the video was recorded in the evening hours. The third video contains a moderately congested segment, and its picture quality is not good. The fourth video has a severely congested segment, and its picture quality is also not good. These test videos are shown in Fig. 4, and the videos are labeled "OD video 1," "OD video 2," "OD video 3," and "OD video 4," respectively.
In addition, a total of four test videos were used for object tracking. The first video does not contain any congested segment, and it was recorded when the conditions were relatively pleasant. The second video contains a congested segment and is very cluttered. The third video does not contain any congested segment, but it was recorded in the evening hours. The fourth video has blurry picture quality owing to rain. These test videos are shown in **Fig. 5**, and the videos are labeled "OT video 1," "OT video 2," "OT video 3," and "OT video 4," respectively.

4.2 Result

4.2.1 Object Detection

Object detection is first performed using Yolov4 prior to tracking objects. To evaluate the performance of object detection, mean average precision (mAP) was used as the performance evaluation index. The mAP is an index that is derived from precision and recall, which are used as the performance indicators for binary classifiers. The mAP can be computed by obtaining the mean of the average precision values, and the AP is then calculated by taking the area under the precision-recall curve. Precision and recall can be calculated using the following formulas.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

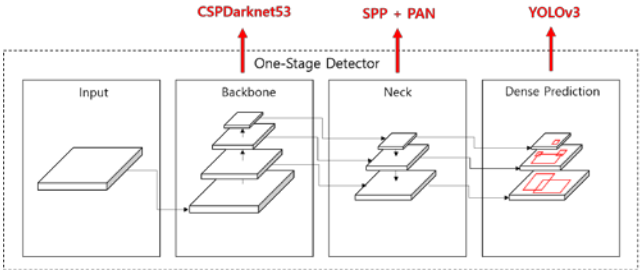


Fig. 2. The overall structure of YOLOv4

Fig. 3. Training data and validation data used in object detection

Table 1. The number of training and validation data used for object detection in CCTV videos

	Trainval video 1	Trainval video 2	Trainval video 3
Training	11268	18460	56550
Validation	2818	4615	14138
Total	14086	23075	70688

Where TP denotes True positive, FP denotes False positive, FN denotes False negative.

In this study, the object detection project performed using Yolov4 achieved 89% mAP. The video images of the object detection result are shown in **Fig. 6**. It can be seen that most of the objects have been detected irrespective of the presence of a congested section and regardless of whether it was day or night. Note that object detection was performed below the red line in the video images.

4.2.2 Object Tracking

Using the DeepSORT model, object tracking was performed based on the result of object detection performed using Yolov4. Note that object tracking was performed below the red line in the video images.

The first video tested does not have any congested segment, and it was recorded when the conditions were relatively pleasant. As shown in **Fig. 7**, most of the vehicles were detected and they were tracked without confusion of the unique IDs.

The second video that was tested contained a congested segment and was very cluttered. As shown in **Fig. 8**, vehicles were detected without any issue, but the vehicles could not be tracked well when they were obscured owing to the severe traffic congestion.

The third video tested did not contain any congested segment, but it was recorded in the evening hours. As shown in **Fig. 9**, most of the vehicles were detected without much issue even though the video images were relatively dark and it was difficult to detect objects in the images. The vehicles were tracked without confusion of the unique IDs.

The last video tested had blurry picture quality owing to rain. As shown in **Fig. 10**, due to the light smearing and blurry picture quality, vehicles could not be tracked even when they were detected correctly.

5. Conclusion and Future Work

In this study, the Yolov4 network was trained on the CCTV videos from three different roads

OD video 1 OD video 2 OD video 3 OD video 4

Fig. 4. Object detection test videos

OT video 1 OD video 2 OT video 3 OT video 4

Fig. 5. Object tracking test videos

OD video 1 OD video 2 OD video 3 OD video 4

Fig. 6. Object detection result

Fig. 7. Object tracking result for the video containing relatively nice condition without any congested segment

Fig. 8. Object tracking result for the video containing relatively nice condition without any congested segment

Fig. 9. Object tracking result for the nighttime video without any congested segment

Fig. 10. Object tracking result for the video with blurry picture quality due to rain

(Sanmakgongdan entrance 3-way intersection, Seoksan bridge 4-way intersection, Deokgye 4-way intersection).

Using the weight file obtained from running object detection, object tracking was performed using the DeepSORT model. The results showed that most of the vehicles could be detected irrespective of the presence of a congested section and whether it was day or night. For object tracking, the vehicles were tracked correctly irrespective of whether it was day or night only when there was no traffic congestion. When the videos contained a congested segment, many of the objects overlapped, making it difficult to track them. It was also difficult to track the vehicles when the video's picture quality was degraded drastically due to rain or other causes, irrespective of whether there was a congested section. In future studies, the tracking model will be enhanced to better track objects even when they overlap, or the video has low image quality.

References

[1] Alonso, J.D., Vidal, R.R., Rotter, A., Muhlenberg, M., "Lane-change decision aid system based on motion-driven vehicle tracking", in *IEEE Transactions on Vehicular Technology*, Vol. 57, No. 5, pp. 2736-2746, 2008.

[2] Tsai, L.W., Hsieh, J.W., Fan, K.C., "Vehicle detection using normalized color and edge map", *IEEE transactions on Image Processing* Vol. 16, No. 3, pp. 850-864, 2007.

[3] Zhang, H., Wu, K., "A vehicle detection algorithm based on three-frame differencing and background subtraction", *Fifth International Symposium on Computational Intelligence and Design*, Vol. 1, 2012.

[4] Unzueta, L., Nieto, M., Cortes, A., Barandiaran, J., Otaegui, O., Sanchez, P., "Adaptive multicue background subtraction for robust vehicle counting and classification", in *IEEE Transactions on Intelligent Transportation Systems*, Vol. 13, No. 2, pp. 527-540, 2012

[5] F. Han, Y. Shan, R. Cekander, H. S. Sawhney, and R. Kumar, "A two-stage approach to people and vehicle detection with hog-based SVM", Proc. of *Performance Metrics for Intelligent Systems 2006 Workshop*, pp. 133-140, 2006.

[6] Chen, C.T., Su, C.Y., Kao, W.C., "An enhanced segmentation on vision-based shadow removal for vehicle detection", *The 2010 International Conference on Green Circuits and Systems*, pp. 679-682, 2010.

[7] Nicolai Wojke, Alex Bewley, and Dietrich Paulus. "Simple online and realtime tracking with a deep association metric", *In 2017 IEEE International Conference on Image Processing*, pp. 3645–3649, 2017.

[8] A. Bewley, G. Zongyuan, F. Ramos, and B. Upcroft, "Simple online and realtime tracking", *In 2017 IEEE International Conference on Image Processing*, pp. 3464–3468, 2016.

[9] Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M., "Yolov4: Optimal speed and accuracy of object detection", *arXiv preprint arXiv:2004.10934*, 2020.

[10] Chien-Yao Wang, Hong-Yuan Mark Liao, Yueh-Hua Wu, Ping-Yang Chen, Jun-Wei Hsieh, and I-Hau Yeh. "CSPNet: A new backbone that can enhance learning capability of cnn", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshop*, 2020.

Media Violence Exposure and Public Confidence in the Police: Examining the Mediating Role of Fear of Crime

Yongil Lee[1], Donghan Lee[1], Dongyeong Maeng[1], Eunhye Kim[1], Jihui Shin[1], EuiGab Hwang[6*]
[1]Department of Criminology, Kyonggi University
[6]Department of Police Administration, Kyonggi University
[flyhwang@gmail.com]
*Corresponding author: EuiGab Hwang

Abstract

Confidence in the police is a driving force for the police to perform law enforcement and crime prevention activities effectively. There exist numerous research articles on confidence in the police, but less attention has been given to the effect of media violence exposure and the mediating role of fear of crime. To fill this void, this study attempted to examine the mediating role of fear of crime in the effect of media violence exposure on confidence in the police by using the Korea National Crime Victimization Survey data collected from 12,620 citizens throughout South Korea. Results revealed the significant effect of media violence exposure directly on confidence in the police, and indirectly via fear of crime even after controlling for major conceptual factors (i.e., direct and indirect victimization, perceived disorder, social cohesion, informal social control). Implications for policies and future research were discussed based on the findings.

***Keywords*:** Media exposure, Fear of crime, Confidence in the police, Expressive perspective, Instrumental perspective, Social integration model, Disorder model, Victimization model, Cultivation theory, Collective efficacy

1. Introduction

The so-called nation's police institute performs most proficiently when it is accepted by the public as a legitimate law enforcement executioner. The police confidence, therefore, stretches two key factors on this research: it is one of the most crucial criteria when scrutinizing police campaigns, and it also becomes a driving force for the police when performing law enforcement activities. There exist numerous research articles on confidence in the police and fear of crime based on several conceptual models, but less attention has been given to the effect of media violence exposure. It also appears that fear of crime has often been recognized as a direct factor of confidence in the police, but not as a mediating factor even though both fear of crime and confidence in the police share common driving factors. To fill this void, it is important to consider media violence exposure as a factor of confidence in the police and also fear of crime as a mediator between conceptual factors and confidence in the police.

2. Literature review

2.1 Conceptual Factors of Fear of crime

Victimization model insists that victimization experience escalates fear of crime. Garofalo (1979) and Lee (2011) imply that crime

victimization is a strong predictor of fear of crime. Indirect victimization has been studied that it contributes to escalate the public's fear of crime even more than direct victimization. Individuals who indirectly involve in criminal activities – by being exposed by the mass media, being around the victims – are frightened by the crime though such activities rarely happen in their everyday life [3, 4, 5]. Crime-related news is an important factor in explaining this escalation of fear of crime [6].

Disorder model and social control model are conceptual cause of fear of crime. As individuals perceive physical and social disorder, it strengthens fear of crime. Inhabitants in a structurally disorganized community are likely to manifest even further intensified fear of crime when they have certain aspects: weak bonding, low willingness for informal social control, and low confidence in the police [7]. Several studies examined this model and found a significant positive relationship between perceived disorder and fear of crime [8, 9, 10, 11].

Sampson et al. (1997) suggested that 'collective efficacy', which is an informal social control based on trustworthy of community residents, can reduce fear of crime. Collective efficacy also measures the level of social integration.

2.2 Factors of Confidence in the police

Instrumental perspective and expressive perspective are main conceptual explanations of confidence in the police. Instrumental perspective on the confidence in the police relies on a perception that the law enforcers predominantly focus on reducing crime and building public safety. The public's confidence in policing refers to their direct or indirect victimization which generates a predominant assessment that the police have poor control over crimes; thus, low confidence in the police [13,14,15].

Expressive perspective on confidence in the police has certain elements to explain confidence in the police: perceived disorder, social cohesion of residence, trust among community members, and moral agreement [16,17]. Jackson and Sunshine (2007) showed that both fear of crime and confidence in the police are affected by public's evaluations of social disorder and cohesion.

2.3 Media Exposure and Confidence in the Police

Cultivation theory understands the mass media as a catalyst that accelerates consumers' fear of crime. Usually, citizens' evaluations of the police are conducted indirectly through the media, which naturally plays an important role in shaping the image of law enforcement. Surette (2015) argued "Because the public has little direct contact with police, the media naturally plays an important role in shaping their perceptions of law enforcement, a role which can be understood through the lens of the cultivation theory". Especially, Gerbner (1970) asserted that there is overflowing sensational news about crime in the mass media, so the more individuals are exposed to these media, the greater fear of their own safety would be felt.

3. Methodology

3.1 Research Design

In this study, the effects of media exposure related to crime as well as experience of direct and indirect crime victimization, perception of incivilities, and collective efficacy on fear of crime and confidence in the police are examined, and analysis of indirect influence of such factors on confidence in the police is conducted by using fear of crime as a mediator.

3.2 Samples

The data of National Crime Victimization survey conducted in 2016 were obtained from the Korean Institute of Criminology. The survey has been recognized to be well established. The national sample of 12,620 citizens were contacted for face to face interviews.

3.3 Measurement

3.3.1. Dependent variable

Confidence in the police was measured by three items. These items showed a high level of reliability (Cronbach alpha = .79).

3.3.2. Mediating variable

Fear of crime was measured with eight items, and the reliability level was significantly high (Cronbach alpha = .931).

3.3.3. Independent variables

The independent variables included media violence exposure, direct and indirect victimization, perceived disorder, social cohesion, informal social control.

The item designed to measure *media violence exposure* was "I often watch news or programs related to crime." For measuring *direct victimization*, participants were asked whether they were victimized in various types of violent crimes. *Indirect victimization* was measured using two items. *Perceived disorder* was measured by six items. The six items showed a high level of reliability (Cronbach alpha = .830). This research measured collective efficacy into social cohesion and informal social control based on the previous studies. *Social cohesion* was measured by four items. These items showed a high level of reliability. (Cronbach alpha = .911) Three items were used to measure *informal social control*, and the reliability was satisfactory (Cronbach alpha = .703).

4. Findings

4.1 Descriptive statistics

Table 2. Descriptive statistics

Variables	Min.	Max.	Mean	SD
Confidence in the police	1	5	3.44	0.68
Fear of crime	1	5	2.10	0.76
Media violence exposure	1	5	3.02	1.04
Direct victimization	0	1	0.03	0.18
Indirect victimization	0	1	0.04	0.20
Perceived disorder	1	5	2.14	0.66
Social cohesion	1	5	2.72	0.95
Informal social control	1	5	3.33	0.77
Sex	0	1	0.53	0.49
Age	14	99	48.95	18.54
Education	1	7	4.22	1.41
Income	1	9	4.11	1.94

Table 2 reports descriptive statistics. For the individual variables, the average age was 48.95 (SD = 18.54), ranging from 14 to 99. There were slightly fewer females (49.9 percent) than males in the sample. The average level of education was 4.22 (SD = 1.41), which corresponds to 'college (less than 4 years)'.

The average pretax income of household was 4.11 (SD = 1.94), which corresponds to the '3 million ~ less than 4 million Korean Won per month'.

For the dependent variable, the average assessment on their own confidence in the police was 3.44 (SD = 0.68), which was slightly above the midpoint.

For the mediating variable, the mean score of fear of crime was 2.10 (SD = 0.76), showing slightly lower level of fear of crime.

For the independent variables, the average level of media violence exposure was 3.02 (SD = 1.04), indicating that More than average of respondents was exposed to crime-related media. And 3% of respondents indicated that they had experienced crime victimization in the past year, 4% indicated that they had acquaintances who experienced crime victimization. The mean score of perceived disorder was 2.14 (SD = 0.66), and social cohesion had 2.72 mean score (SD = 0.95). Last, the average informal social control was 3.33 (SD = 0.77) which was a bit higher than the mid-point.

4.2 Bivariate correlation among the variables

The results indicate that the factors are significantly correlated to each other. In particular, confidence in the police showed the highest correlation with informal social control ($r = 0.322$, $p < .01$). Confidence in the police also showed a positive relationship with media violence exposure ($r = 0.083$, $p < .01$), but it showed a negative relationship with fear of crime ($r = -0.062$, $p < .01$). For the fear of crime as a mediator, fear of crime showed the highest correlation with perceived disorder ($r = 0.349$, $p < .01$). And only age showed a negative relationship with fear of crime.

4.3 Structural equation modeling

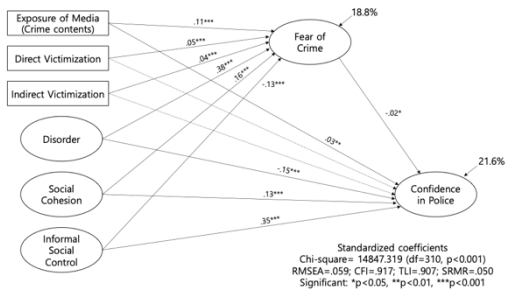


Fig. 1. Determinants of public confidence in police

The model fit showed acceptable approximate-fit indices (RMSEA = 0.059; CFI = 0.917, TLI = 0.907, and SRMR = 0.050). The chi-square test

showed significance (= 14847.319, df = 310, p < .001). For confidence in the police, all of the independent variables showed significant relationships except experience of victimization. It was confirmed that all of the independent variables used in this analysis had significant effects on fear of crime. Specifically, Media exposure, victimization experience, disorder, and social cohesion was significantly positively linked to the fear of crime, and informal social control had a negative relationship with fear of crime. Interestingly, as social cohesion increased, fear of crime also increased. This result is in contrast to the previous research results. Eventually, as a mediating variable, fear of crime affected negatively on confidence in the police in a significant way. The result confirms that fear of crime (MV) and perception of the police (DV) have a negative association.

Table 3. Structural effects via fear of crime

Path	Standardized coef.
Media violence exposure → Fear of crime → Confidence in the police	-0.003*
Direct victimization → Fear of crime → Confidence in the police	-0.001*
Indirect victimization → Fear of crime → Confidence in the police	-0.001*
Disorder → Fear of crime → Confidence in the police	-0.010*
Social cohesion → Fear of crime → Confidence in the police	-0.004*
Informal social control → Fear of crime → Confidence in the police	0.004*

Table 3 shows how fear of crime (MV) mediates major factors and confidence in the police (DV). Media violence exposure, direct victimization, indirect victimization, perceived disorder, and social cohesion affected negatively confidence in the police via fear of crime. In particular, media exposure increases fear of crime, which eventually lessens confidence in the police. And only informal social control has a positive effect on the confidence in the police via fear of crime.

5. Conclusion

Confidence in the police is a critical element of policing, as the police are able to perform their duties when the public view them as a legitimate law enforcement and crime prevention actors. This study gave attention to media violence exposure as a factor of both fear of crime and confidence in the police, and fear of crime as a mediator. Findings supported our structural hypothesis. Results showed the significant effect of media violence exposure directly on confidence in the police, and indirectly via fear of crime even after controlling for major conceptual factors (i.e., direct and indirect victimization, perceived disorder, social cohesion, informal social control). Exposer to criminal activities through the media marginalized confidence in the police, and conceptual factors were linked to confidence in the police via fear of crime.

Results imply that the role of media has so far not been recognized enough as a driving factor of both fear of crime and confidence in the police. Additionally, policies to enhance public confidence in the police may need to give more attention to fear of crime as it plays a mediating role between conceptual factors and confidence in the police.

References

[1] J. Garofalo, "Victimization and fear of crime," *Journal of Research in Crime and Deliquency*, Vol 16, No.1, pp. 8-97, 1979

[2] J. Lee, "A study on the interaction effect between victimization and community characteristics on fear of crime," Journal of Korean Criminal Psychology, Vol 7, No.3, pp. 231-250, 2011

[3] W. G. Skogan, M. G. Maxfield, "Coping with crime," *Sage publication*, 1981

[4] J. Choi et al, "Influence of Media on the Fear of Crime of Martial Migrant Woman : An Empirical Study focused on Cultivation theory and Vulnerability Hypothesis," *The Korea Police Science Journal*, Vol 16, No. 6, pp. 172-192, 2014

[5] A. D. Biederman, A. J. Reiss, "On exploring the 'dark figure' of crime," *The Annals of the American Academy of Political and Social Sciences*, 374, pp. 1-15, 1967

[6] Noh et al, "The effect of criminal victimization and fear of crime on satisfaction of life in youth,"*Korean Journal of Victimology*, Vol 21, No.1, 2012

[7] R. B. Taylor, M. Hale, "Testing alternative models of fear of crime," *Journal of Criminal Law and Criminology,* 77, pp. 151-189, 1986

[8] T. W. Franklin, C. A. Franklin, N. E. Fearn, "A Multilevel Analysis of the Vulnerability, Disorder, and Social Integration Models of Fear of Crime," *Social Justice Research,* Vol 21, No.2, pp. 204-227, 2008

[9] G. J. DeLone, "Public housing and the fear of crime," *Journal of Criminal Justice,* Vol 26, No.2, pp. 115-125, 2008
[10] W. J. Wilson, "When Work Disappears: The World of the New Urban Poor, Knopf, 1996
[11] A. Villarreal, B. Silva, "Social cohesion, criminal victimization and perceived risk of crime in Brazilian neighborhoods," *Social Forces,* Vol 84, No.3, pp. 1725-1753, 2006
[12] R. J. Sampson, S. Raudenbush, F. Earls, "Neighborhoods and violent crime: A multilevel study of collective efficacy," *Science,* Vol 277, No.5328, pp. 918-924, 1997
[13] J. Jackson, B. Braford, K. Hohl, S. Farrall, "Does the fear of crime erode public confidence in policing?," *Policing: A Journal of Policy and Practice*, Vol. 3, No.1, pp. 100-111, 2009
[14] K. M. Donovan, I. C. F. Klahm, "The role of entertainment media in perceptions of police use of force," *Criminal Justice and Behavior,* Vol 42, No.12, pp.1261-1281, 2015
[15] J. Jackson, J. Sunshine, "Public confidence in policing: a neo-Durkheimian perspective," *British Journal of Criminology,* Vol 47, No.2, pp. 214-233, 2007
[16] J. Jackson, J. Sunshine, "Public confidence in policing: A Neo-Durkheimian Perspective," *The British Journal of Criminology,* Vol 47, No.2, pp. 214-233, 2006
[17] J. Jackson, B. Bradford, "Crime, policing and social order: on the expressive nature of public confidence in policing," *The British Journal of Sociology,* Vol 60, No.3, pp. 494-521, 2009
[18] R. Surette, "Performance crime and justice," *Current Issues in Criminal Justice,* Vol 27, No.2, pp. 195-216, 2015
[19] G. Gerbner, "Cultural indicators: The case of violence in television drama," *Annals of the American Association of Political and Social Science,"* Vol 338, pp. 69-81, 1970

Table. 1. Confirmatory factor analysis

Variables	Items	Loadings	Cronbach' a
Dependent variable			
Confidence in the police	1. I think the police in our neighborhood are doing well in patrolling our area. 2. I think the police in our neighborhood will come to the crime scene immediately, if a crime happens. 3. I think the police in our neighborhood will arrest the offenders of crime, if we report the crime incidents.	0.797 0.815 0.833	.790
Mediating variable			
Fear of crime	1. I am afraid that someone might steal my money or property without my knowledge. 2. I am afraid that someone might take my money or property by force. 3. I am afraid that someone might assault and hurt me. 4. I am afraid that someone might take my property by fraud 5. I am afraid that someone might sexually harass or assault me. 6. I am afraid that someone might vandalize my property. 7. I am afraid that someone might break into my house. 8. I am afraid that someone might stalk me or harass me by phone.	0.827 0.940 0.825 0.794 0.749 0.831 0.803 0.811	.931
Independent variable			
Media violence exposure	1. I often watch news or TV shows about crime	0.506	
Direct victimization	1. In 2016, have you ever suffered from property laws by frauds and swindles? 2. In 2016, have you ever had your property stolen or taken unlawfully, or nearly stolen or taken unlawfully? 3. Aside from the incidents you have mentioned already, has anyone broken into or attempted to break into your house cutting through the security grill or lock or opening a window or a door in 2016? 4. Aside from the incidents you have mentioned already, has anyone deliberately damaged or vandalized any of your or a family member's property in 2016? 5. Aside from the incidents you have mentioned already, have you ever been assaulted, threatened, or robbed, or nearly assaulted, threatened, or robbed in 2016? 6. Aside from the incidents you have mentioned already, have you ever been assaulted, threatened, or almost assaulted or threatened in any of the following ways in 2016? 7. People tend to underestimate the damage inflicted by acquaintances. Aside from the incidents you have mentioned already, have you ever been robbed, assaulted, threatened, or harassed, or nearly robbed, assaulted, threatened, or harassed by any one of the following in 2016? 8. Aside from the incidents you have mentioned already, have you ever or almost experienced unwanted sexual contact by any one of the following in 2016?	0.604 0.571 0.697 0.264 0.821 0.908 0.875 0.757	
Indirect victimization	Has anyone close to you (family, relatives, friends, colleagues) experienced the following in 2016? 1. Physical injury 2. Property damage	0.571 0.737	
Perceived disorder	1. The streets were dirty and littered with trash. 2. There were a lot of blind spot. 3. There were a number of abandoned cars or blighted properties. 4. A large number of people violated simple traffic laws. 5. There were many indecent teenagers roaming around the neighborhood in groups. 6. People were frequently seen arguing loudly or fighting.	0.768 0.730 0.641 0.717 0.728 0.745	.830
Social cohesion	1. My neighbors knew each other well. 2. I frequently talked about things around the neighborhood with my neighbors. 3. My neighbors were willing to help each other. 4. My neighbors actively participated in events and meetings in the neighborhood.	0.862 0.886 0.868 0.842	.911
Informal social control	1. My neighbors would not hesitate to offer help if a neighborhood kid is being harassed by kids from out of town. 2. My neighbors would report a crime to the police if they witness one. 3. My neighbors would volunteer as neighborhood watchers if it is necessary to prevent crime.	0.759 0.831 0.487	.703

An Active Contents-Bigdata Engineering Framework for Intelligent-led Crime Prevention and Prediction

Kyunghee Sun[1], Jeong-Hyun Chang, Hyun Yoo, Kyoung-Sook Kim[1], Hyun Ahn[1] and Kwanghoon Pio Kim[1*]
[1] Contents Convergence Software Research Institute
Kyonggi University, Republic of Korea
[e-mail: {phamdinhlam, hahn, khmjmc, kwang}@kgu.ac.kr]
*Corresponding Author: Kwanghoon Pio Kim

Abstract

This paper presents a conceptual initiation of active contents-Bigdata engineering and its practical application in the intelligent-led policing and crime prevention activities. The concept of active Bigdata is concretized into an engineering framework for collecting, preprocessing, exploring, mining and deep-learning efficiently and effectively such unstructured contents-Bigdata like texting data, images, video clips from social network services and video clips from CCTV. At Kyonggi Univ., we have established a new research institute focusing on fulfilling the full functionality from the very early stage of building a novel architectural framework for supporting the active contents-Bigdata engineering activities to the very late stage of applying the implemented framework to a very practical and impeccable solution of intelligent-led policing and crime prevention as a separate discipline as well as a convergence discipline of computer science and criminal justice and psychology. The new research institute organizes a conference session of "Bigdata and AI Technologies for Crime Prevention and Prediction" for holding an open discussion as well as a knowledge exchanging opportunity on the research issues and topics. In this paper as a positioning paper, we describe the research goal and the functional scope of the active Bigdata engineering framework and raise the open issues for the application of the framework to the convergence discipline of the intelligent-led policing and crime prevention and prediction.

***Keywords*:** Bigdata, Active Contents-Bigdata, Intelligent Bigdata Analytics, Contents Tagging and Activating, Bigdata Platform, Active Contents Convergence, Active Contents Management, Intelligent-led Policing, Crime Prevention and Prediction, Machine Learning, Deep Learning

1. Introduction

Today's societies and lives are characterized by two phenomena of data-floodings[1,2,3,4], video clippings of data-floodings and criminal scenes of data-floodings, on the cyber world as well as the physical world. At the height of data-flooding era, we all, as data scientists and criminal psychologists, are trying to detect and analyze a variety of criminal scenes and situations by using a huge amount of (structured and unstructured) data-stores[5] obtained from these data-floodings phenomena in order to maintain the public safety[6] and security[7]. Moreover, we all are threatened from a lot of social criminal patterns[8] that we haven't seen and experienced before at all. Moreover, in terms of the victims' scope, almost all the generations of human and

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(grant number 2020R1A6A1A03040583).

even their animal companions are becoming the subjects and the targets of crimes at almost everywhere. Therefore, we need a kind of silver-bullets for timely and completely controlling these criminal societies, and we all definitely believe that using these data-floodings phenomena must be an impeccable and effective silver-bullet, and throuth which we are able to dig into much more effective preparations and solutions for coping with such unexpected criminal societies[6,7] and for keeping and maintaining the public safety and security, as well.

As one of the trials making the scientific and engineering preparations, at Kyonggi Univ., we have organized a new research group looking for a convergence approach of computer science, data science and criminal justice disciplines to make a wide variety of Bigdata-engineering solutions for supporting Bigdata-driven and intelligent-led criminal prevention and prediction activities and eventually contribute to realize the public safety and security society being liberated as well as being set free from the unexpected threatening and attacking criminals.

In this paper as a positioning paper, we introduce the research goal and scope of the convergence research group through setting up an engineering framework that is to concretize the novel concept of active contents-Bigdata that are formatted in XML-tagging standards and obtained from huge amounts of unstructured data-stores[9,10], such as texting data-stores and CCTV video-clip data-stores, and others. The engineering framework is for fulfilling the full functionality from the very early stage of building a novel architectural framework for supporting the active contents-Bigdata engineering activities to the very late stage of applying the implemented framework to a very practical and impeccable solution of intelligent-led crime prevention and prediction as a separate discipline as well as a convergence discipline of computer science and criminal justice and psychology.

In terms of organizing the paper, we describe the research goals and scope in the next section. And through additional three consecutive sections, we introduce the major research themes and their functional components and architectural details. Finally, we finalize the position paper by summarizing the research topics with the future works to be performed in the convergence research group that is the firstly organized research trial for cooperatively working together by a convergence professional group of computer scientist, data scientist and criminal psychologist in South Korea.

2. Research Goals and Scope

The convergence research group has outlined the goals and technological scope for the engineering framework to be described in this positioning paper. The outlined is defined into two aspects of the engineering research issue and the application research issue. The former is about accomplishing the concretization of the active contents-Bigdata, the latter is about applying the framework to the crime prevention and prediction activities with respect to the theme of "***Stopping crime before it starts!***" that has been stated as an ideal and eventual goal in the arena of intelligent-led policing and crime justice and psychology discipline and industry.

The scope of the engineering aspect is defined in three technical requisites, which are closely related with the so-called trinity of the Bigdata engineering techiques, acquisition, platform and analysis: 1) unstructured contents-Bigdata tagging and activating requisites, 2) active contents-Bigdata platform engineering requisites, and 3) active contents-Bigdata analysis and deep learning requisites. The scope of the application aspect is also defined by applying the implemented framework to help the eradicating activities of the first three out of the top 4-vice, such as sexual violence, domestic violence, school violence and adulterated food, that are regulated by the ministry of justice of Korean government.

Conclusively speaking, the new concept of active contents-Bigdata engineering proposed in this paper can be concretized as the active contents-Bigdata engineering framework that is able to properly handle, preprocess, acquisition, and analyze such unstructured video and social media contents-Bigdata data-stores. As a consequence of the newly defined concept, the framework ought to be fortified with the three mighty achievements of technological supports: the tagging and activating mechanisms, the platform engineering and deploying architectures and the analyzing[11,12] and

deep-learning models[13,14] and services[15,16], and it is named as active contents-BIGdata engineering framework. Also, through long-term and intensive supports from the national research foundation of Korea, the challenging topics and their related technologeis of the active contents-Bigdata engineering framework will be developed and implemented, and also the application of the framework will be successfully planned and fulfilled in the three-vice of social crimes and we ought to be able to prove the crime prevention effects through the implemented framework. Finally, it is possible to intelligently integrate and optimize the crime prevention and prediction functions throughout the city to be deployed and testified, and we are able to ultimately provide a safe quality of life to the public and the people from various social risk factors and threatening criminals.

3. Active Contents-Bigdata Engineering Framework

In this section, we describe the detailed functional components of the active contents-Bigdata engineering framework proposed in this paper. As stated in the goals of the research group, the framework is theoretically based upon the concept of active contents-Bigdata and the conceptual operation of the contents-activation. We start introducing these concepts and finishing their detailed implementations. In particular, the meaning of the contents-activation implies, which is a newly invented operation by the research group, the operation transforming from the original contents holding the passive property themselves to the active contents embedding the active property like analysis functions and their analyzed results formatted in the standardized XML-tags. **Fig. 1** illustrates the major functional components of the engineering framework, which is consisting of the active Bigdata acquisition engineering technology, the active Bigdata platform engineering technology and the active Bigdata analysis engineering technology. The following subsections describe the functional details of these three technologies.

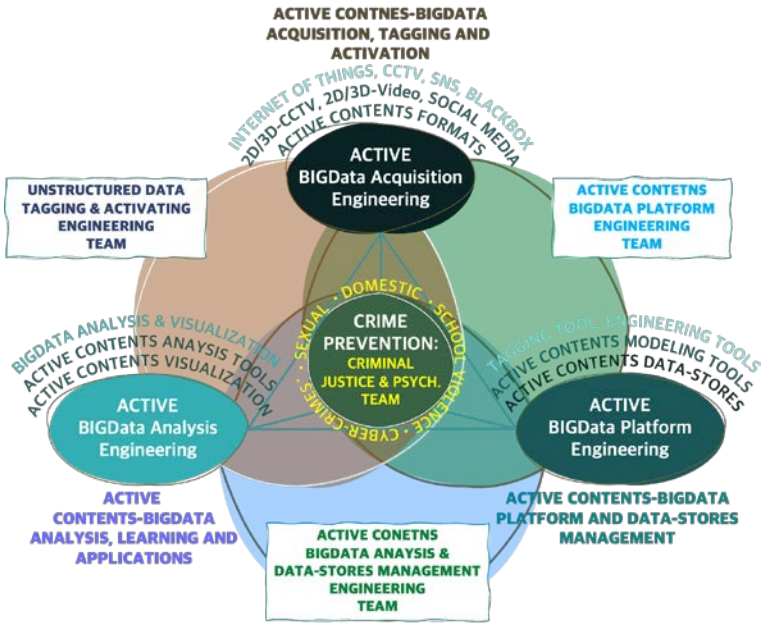


Fig. 1. Major Functional Components of the Engineering Framework

3.1 acBigdata Acquisition

The active contents-Bigdata, abbreviated as acBigdata, acquisition engineering technology consists of the tagging mechanisms and the packaging methodologies for supporting the active contents-Bigdata acquisition activities that are eventually building a data-store in a standardized XML-format. One of the tagging mechanisms is to define an XML-based tag schema structure and construct a bunch of active

contents-objects by using the XML-based tagging schema and format. Especially, we will define a series of the structural types formatting such active contents objects, like single modality active contents-objects, dual-modality active contents-objects with properties, and multi-modality active contents-objects. The packaging methodology is about defining an active contents package structure that is made up of active contents-object model, active contents aggregation model and active contents assets with holding their operations and tagging information. Additionally, the packaging methodologies need to complement the block-chain technology for providing the information security and governance of execution history on each of the active contents-objects. **Fig. 2** shows the tagging mechanisms with XML schema structures.

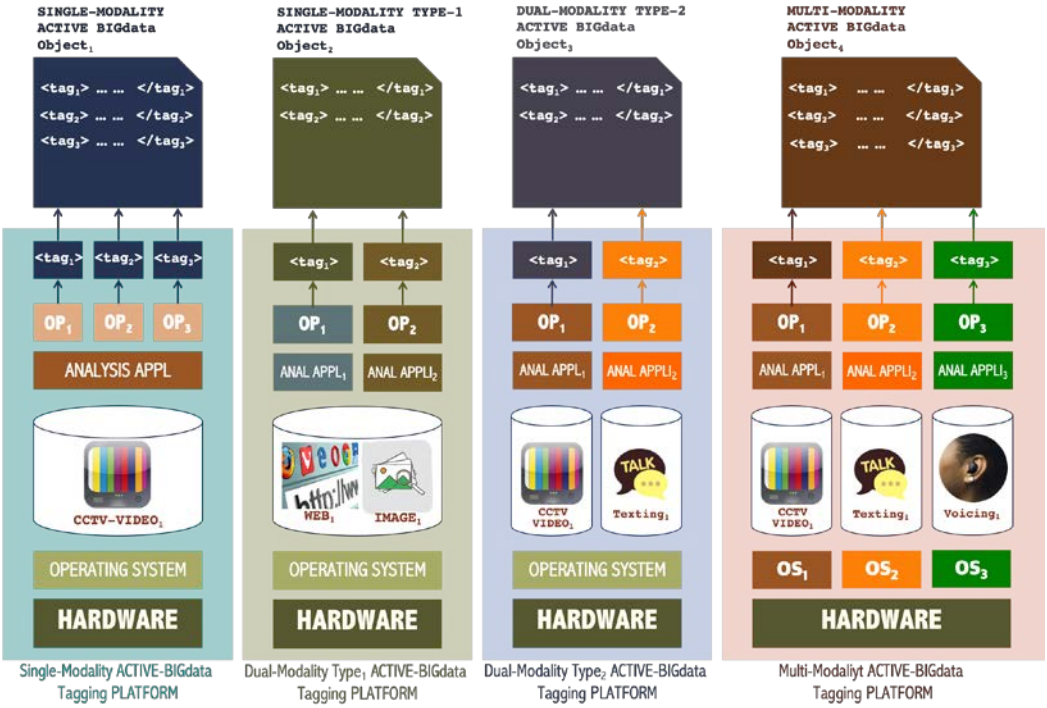


Fig. 2. Tagging Mechanisms: XML-based Tag Schema and Formats

3.2 acBigdata Platforms

After successfully setting up the active contents-Bigdata acquisition engineering mechanisms and methodologies, we need to build a couple of platforms and their architectures for generating, operating, maintaining, and monitoring those active contents-Bigdata objects and their data-stores. These platforms are the core components of the active contents-Bigdata computing environments. The first platform to be implemented in the research group ought to be the active contents-objects generating and archiving platform equipped by the block-chain solution for tracking the execution histories and records. Another core components of the platforms is to provide an operational environment for storing, updating, managing, monitoring the active contents-object data-stores organized with their tagging information that are divided into three units of tagging-elements, such as object, operation and situation, captured from the original source datasets like texting Bigdata, images Bigdata and CCTV video clip Bigdata.

3.3 acBigdata Analysis

The final essential technology for performing the active contents-Bigdata engineering activities is obviously to be related with analyzing the active contents-object data-stores and applying them to the crime prevention and prediction works,

which are collected as well as managed by the acBigdata acquisition and acBigdata platform engineering tools developed by both of the previous research groups, respectively. There are five analysis research trials to be conducted by the research group taking charge of this analysis and application technology. The first trial is about providing the analyzing platform and its crime prevention and prediction services based upon the active contents-object data-stores and their engineering support tools. The second trial is about developing the object detecting, searching and tracing engines over the active contents-object data-stores. Especially, the research group wants to challenge to develop an analysis tool that is supporting the object-detection and moving-object tracing functionality on 2D/3D video clips. The third trial is about the prediction technology based upon the machine learning and deep learning techniques. Through these prediction models and their services, we definitely are able to detect and even predict a variety of criminal objects and situations on those CCTV video clips. As the deep learning models and engines, we can use a series of advanced neural network models and their learning engines, such as generative adversarial networks, autoencoder, convolutional neural networks, recurrent neural networks, long short-term memory neural networks, and others. The fourth trial is about analyzing a kind of emotional status of criminals through their voices and layered voice tones. In this research trial, we also use those deep learning model and engines, too, listed in the previous section. Finally, the fifth trial is about detecting and extracting video-clips through automatic operations, recommendation systems and mining techniques from the active contents-object data-stores.

3.4 acBigdata Applications

As the eventual application of the active contents-Bigdata engineering framework, the research institute will develop a novel crime prevention approach that is able to provide a series of crime prevention services impeccably dealing with the Korea's top-4 violence types as well as almost all the cyber-crime types through successfully deploying the capability-maturity model for crime prevention and prediction illustrated in **Fig. 3**.

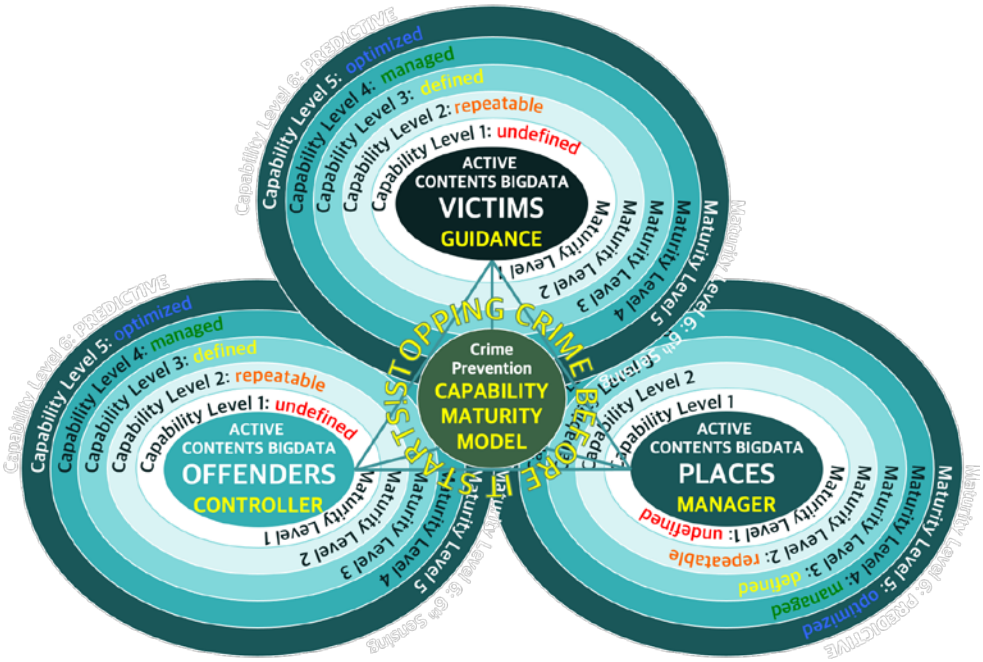


Fig. 3. The Capability-Maturity Model for Crime Prevention and Prediction

The capability-maturity model is based upon the theory of situational crime prevention that is a well-known criminology theory in the criminal justice discipline. It consists of three situational factors of crime-occurrence such as offenders, victims and places where the crime is committed. Note that the controller, guidance and the manager are able to restrain the offenders, guide the victims and fortify the places, respectively, against crimes. Basically, the situational crime prevention theory sets forth as a premise that any crime can be prevented if any of the three situational factors is forbade by any of the corresponding controller, guidance and the manager. Each of the situational factors has five-level of technological capability and five-level of technological maturity, such as undefined, repeatable, defined, managed and optimized levels. Each level of capability implies the degree of how much capable organizations (or public district) is adopting the corresponding technology, whereas each level of maturity implies the degree of how much the corresponding technology is matured in terms of its usability and adoptability. Based upon the capability-maturity model, we can evaluate the level of crime prevention and prediction capability of organizations (or public districts). Due to the page limitation, we won't describe the functional details of the model's components in this paper. We strongly believe that this model ought to be a focal remedy for realizing the intelligent-led policing and crime prevention and prediction strategies in near future.

4. Conclusions

In this paper, we have introduced a novel research framework to be performed by the newly established research institute, which is named as the active contents-Bigdata engineering framework based upon the Bigdata mining and deep-learning technologies. Also, we have a plan to apply this framework to realize so-called the intelligent-led policing and crime prevention and prediction services. Additionally, we will develop a deployment model of this framework, which is named as the capability-maturity model for crime prevention and prediction. As one of the near future works, we will construct an active contents-Bigdata archiving center providing almost all separate professional and situational data-stores in the field of criminal justice and psychology.

References

[1] Hyun Ahn, Kwanghoon Pio Kim, "Formal approach for discovering work transference networks from workflow logs", *Information Sciences*, Vol. 515, pp. 1-25, 2020.

[2] Kwanghoon Pio Kim, "An XPDL-Based Workflow Control-Structure and Data-Sequence Analyzer", *KSII Transactions on Internet and Information Systems*, Vol. 13, No. 3, pp. 1702-1721, 2019.

[3] Dinh-Lam Pham, Hyun Ahn, Kwanghoon Pio Kim, "Discovering Redo-Activities and Performers' Involvements from XES-Formatted Workflow Process Enactment Event Logs", *KSII Transactions on Internet and Information Systems*, Vol. 13, No. 8, pp. 4108-4122, 2019.

[4] Muhammad Habib ur Rehman, Ibrar Yaqoob, Khaled Salah, Muhammad Imran, Prem Prakash Jayaraman, Charith Perera, "The role of big data analytics in industrial Internet of Things", *Future Generation Computer Systems*, Vol. 99, pp. 247-259, 2019.

[5] Blau, P. M., "Structural contexts of opportunities", *University of Chicago Press*, 1994.

[6] Phillips NE, Levy BL, Sampson RJ, Small ML, Wang RQ, "The Social Integration of American Cities: Network Measures of Connectedness Based on Everyday Mobility Across Neighborhoods.", *Sociological Methods & Research*, July 2019. doi:10.1177/0049124119852386

[7] Sampson, R.J., Levy, B.L., "Beyond Residential Segregation: Mobility-Based Connectedness and Rates of Violence in Large Cities", *Race and social problems*, Vol. 12, pp.77–86, 2020. doi.org/10.1007/s12552-019-09273-0

[8] Bora, N., Chang, Y. H., & Maheswaran, R., "Mobility patterns and user dynamics in racially segregated geographies of US cities", in *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction, Springer International Publishing Switzerland 2014*, pp. 11-18. 2014.

[9] Murad Khan, Muhammad Babar, Syed Hassan Ahmed, Sayed Chhattan Shah, Kijun Han, "Smart city designing and planning based on big data analytics",

Sustainable Cities and Society, Vol. 35, pp. 271-279, 2017.

[10] Wissem Inoubli, Sabeur Aridhi, Haithem Mezni, Mondher Maddouri, Engelbert Mephu Nguifo, "An experimental survey on bigdata frameworks", *Future Generation Computer Systems*, Vol. 86, pp. 546-564, 2018.

[11] D. E. Rumelhart, G.E. Hinton, R.J. Williams, "Learning Representations by Back-propagating Errors," *Nature*, Vol. 323, pp. 533-536, 1986.

[12] S. J. Russell and P. Norvig, "Artificial intelligence: a modern approach. Malaysia:" *Pearson Education Limited*, 2016.

[13] K. He, X. Zhang et. al., "Deep Residual Learning for Image Recognition," *In Proc. of CVPR*, pp. 770-778, 2016.

[14] I. J. Goodfellow, D. Erhan et. al., "Challenges in representation learning: A report on three machine learning contests", in *Proc. of Neural Information Processing*, pp. 117-124, 2013.

[15] T. Chen, H. R. Tsai, "Application of Industrial Engineering Concepts and Techniques to Ambient Intelligence: A Case Study", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 9, No. 2, pp. 215-223, 2018.

[16] K. Dhir, A. Chhabra, "Automated Employee Evaluation using Fuzzy and Neural Network Synergism through IoT Assistance", *Personal and Ubiquitous Computing*, Vol. 23, No. 1, pp. 43-52, 2019.

Predicting Next Workflow-Activities Using LSTM-based Deep-Learning Neural Networks

Youngin Park[1], Jaeyoung Yun and Kwanghoon Pio Kim[1*]
[1] Contents Convergence Software Research Institute
School of Computer Science and Engineering, Kyonggi University, Republic of Korea
[e-mail: {youngin1992, lastmore9, kwang}@kgu.ac.kr]
*Corresponding Author: Kwanghoon Pio Kim

Abstract

This paper proposes a LSTM-based neural network model for predicting next workflow-activities in modeling and refining workflow processes as fulfilling a part of the redesigning and reengineering phase. Especially, it is very important to prepare the higher quality of input features as well as the higher volume of input features as much as possible in order to improve the correctness of predicting next workflow-activities after performing the training phase of the deep-learning approach. Therefore, in this paper, we present the detailed description of preprocessing, training and predicting phases in the deep learning approach that is used to construct the LSTM-based neural network model. Finally, we carry out an operational experiment by applying the proposed neural network model to the real dataset of the Review Example Large workflow process event logs that are formatted in the IEEE-XES standard and that are containing 236,360 event logs recorded from the execution of 10,000 traces of workflow process instances.

***Keywords*:** Deep Learning, Recurrent Neural Network, LSTM (Long Short-Term Memory) Neural Network, Workflow Process Redesigning and Reengineering, Information Control Nets, Workflow Process Mining, Next Workflow-Activity Prediction

1. Introduction

Each of workflow and/or business processes is to model a temporal sequence of procedural work-activities deployed and run on process-aware enterprises. Once, a workflow process model is defined and deployed onto the corresponding enterprise, then it implies for its life-cycle to be initiated, maintained and managed, and it is necessary for its redesigning and reengineering to be going on the whole of the life-cycle, as well. In another words, the defined and deployed workflow process model must be redesigned and reengineered continuously and redeployed again. These redesigning and reengineering activities have been systematically supported from exploring, analyzing and monitoring process-aware enterprise bigdata data-stores, such as the workflow and business process mining techniques [1][2][3] and/or the workflow and business process deep learning techniques [4][5][6][7][8].

As a workflow and business process remodeling approach, we have conceived a concept of predictive workflow process modeling method that is based upon the primitive operation of next workflow-activity prediction supported by a LSTM(Long short-term memory)-based [4] deep learning neural network model in this paper. To develop the primitive operation, first of all, we need a well-organized ground-truth dataset containing the event logs and traces recorded

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(grant number 2020R1A6A1A03040583).

from executing workcases and instances of the corresponding workflow process model, so as to be used not only for training the neural network model but also for testing and verifying the proposed deep learning neural network model. Thereafter, we devise a series of algorithms dealing with the three steps, such as preprocessing, training and predicting steps of the deep learning approach.

In terms of organizing this paper, we describe the literture survey results that are related with the topic of so-called predictive monitoring issue, like next workflow-activities, next workflow-performers, remaining time prediction, predictive workload and others, through developing their deep learning neural network models. In the next consecutive sections, we state the detailed description of the proposed primitive operation of the next workflow-activity prediction supported by the LSTM-based deep learning neural network model and service developed in this paper. Finally, we show the experimental results obtained from the application of the propsed deep learning neural network model to a real-dataset of event logs and traces recorded from executing 10,000 instances of the Review Example Large workflow process model.

2. Related Works

The research challenge in this paper is directly related to the predictive monitoring issue and the predictive redesigning, reengieering and remodeling issue, which is hot-issued recently in the workflow process literature. These issues need to be supported by the deep learning neural network architectures and techniques, definitely, as their technological backgrounds. One of the typical deep learning techniques that are fitted very well to the predictive monitoring and remodeling issues ought to be the LSTM-based deep learning neural network architectures and models. The LSTM is a variant of the recurrent neural network (RNN) architectures. That is, the LSTM architecture is able to solve the vanishing gradient problem of the classical RNN deep-learning neural network architectures. Therefore, it is one of the well-known and popular recurrent neural network architectures, which is able to properly handle the prediction works with such time-series datasets, just like workflow process enactment event logs and traces datasets.

In recent, there are many research works [5][6][7][8] that adopt these deep-learning neural network architectures and approaches in the workflow and business process management literature. As a typical trial, J. Evermann, et al. [5] proposed a deep learning approach providing a predictive business process behaviour functionality at run-time. N. Tax, et al. [6] introduced at first the concept of predictive business process monitoring functionality with using LSTM neural network architectures. Another functionality in the predictive business process monitoring concept is the remaining time prediction of business process instances. And, the LSTM neural network models supporting this functionality was firstly delivered by N. Navarin, et al. [7]. Also, S. Ham, et al. [8] carried recently out a LSTM-based deep learning experiment for predicting business process instances' remaining times which is especially featured in activity-centric normalization techniques.

Summarily, in this paper we try to propose a novel concept of the predictive workflow process remodeling that is theoretically and technically supported by the primitive operation of the next workflow-activity prediction with LSTM-based deep learning nerual network architectures. Through the primitive operation proposed in the paper, the novel concept is possibly concretized theoretically and systematically, as well.

3. Next Workflow-Activities Prediction Approach

In this section, we describe the details of the next workflow-activity prediction approach that is based upon the LSTM-based deep learning neural network architectures. The approach consists of three phases such as preprocessing, training and predicting phases as shown in **Fig. 1**. Firstly, we have to refine the XES-formatted event log datasets and produce the time-series datasets of temporal workcases and traces through the preprocessing phase. Secondly, we take a series of training operations for the LSTM-based deep learning neural network model to deeply learn by using the preprocessed datasets of temporal workcases and traces. Thirdly, we test the trained LSTM-based deep learning neural network model to verify how much correctly and exactly it can predict the next workflow-

activities associated with the corresponding workflow process instances (workcases).

3.1 Preprocessing: Time-Series Datasets

The original datasets that were recorded from executing workflow process instances have to be preprocessed to be used for the deep learning prediction models. We devise a new conceptual format of temporal workcases that is used for forming a time-series dataset for the deep learning neural network model. In this preprocessing phase, all the event logs associated with a specific workflow process instance is transformed into a corresponding temporal workcase according to the temporal order of their time-stamps. Note that we won't introduce the detailed description of the temporal workcase concept due to the page limitation. However, we would put emphasis that all those temporal workcases are forming a time-series bigdata-store to be used not only for training the deep learning neural network model as input features, but also for verifying and predicting the next workflow-activities as output features.

In terms of the workflow process instances' event log formats, the workflow and business process literature has been released several standardized formats, such as XWELL [9], BPAF [10], IEEE-XES [11]. These formats have a similar structure of XML-schema. We assume that the event logs of the original data-store are formatted in the IEEE-XES standardized XML- schema as shown in **Fig. 1**.

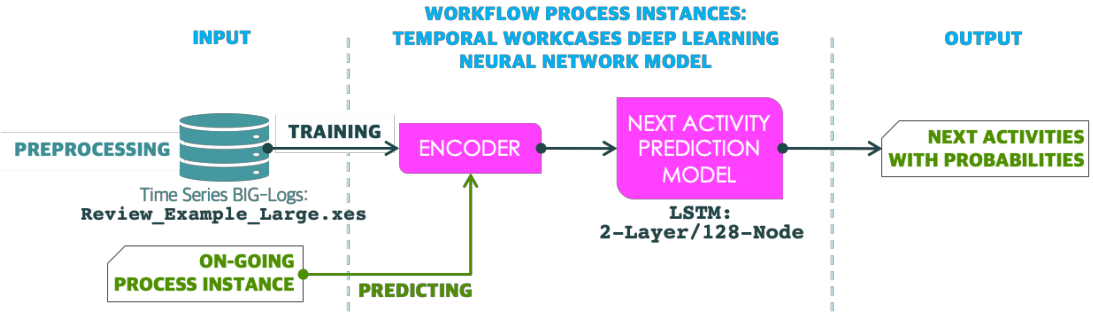


Fig. 1. The Next Workflow-Activity Prediction Approach

3.2 Training: the LSTM-DL model

In the training phase, the preprocessed temporal workcases bigdata-store is used for training the deep learning neural network model that is able to predict the next workflow-activities with their probabilities of each workflow process instance. The unit of an input feature is the activity identifier belonged to each event log in a temporal workcase of the corresponding workflow process instance. The output feature is the possible next workflow-activities with their probabilities. Consequently, the deep learning neural network model developed in this paper is constructed in 2-hidden layers with 128 internal nodes of each as shown in **Fig. 1**. In order to learn the next workflow-activities from the temporal workcase bigdata-store and verify the results of their predictions, we tried to repeatedly train the model at least 200-epoch by using Google's Tensorflow deep learning neural network framework used to make the functional codings with Python programming language.

3.3 Predicting: Next Workflow-Activities

In the predicting phase, we try to predict a series of the next workflow-activities forming an expected temporal workcase by applying to the trained deep learning neural network model. The trained neural network model eventually is able to predict next workflow-activities with probabilities. When an activity-id in a specific temporal workcase comes to the neural network model as input feature, then the model predicts the next workflow-activities of the input activity-id. These predictions are starting from the "START" activity-id and ending to the "END' activity-id, continuously. During fulfilling each of these predictions, it ought to be choosing a workflow-activity as the next one out of the candidate workflow-activities that are outputs of the prediction as output features. For doing these choices, we set the predicted probability as a threshold to make the candidate workflow-activities. Note that the expected temporal workcase generated from the series of the next

workflow-activity predictions must be same to at least one of the possible execution sequences of workflow-activities, which imply all the possible control-paths of the corresponding workflow process model.

4. Operational Experiment

As an operational verification, we carry out an experiment on a dataset of the "Review Example Large" workflow process instances' event logs, which is available in the "4TU center for research data" open data-stores. The dataset is formatted in the IEEE-XES and synthetic dataset for workflow process mining activities. It also contains 10,000 traces (workflow process instances) organized by 236,360 event logs (execution histories of 16-activity's workitems). The corresponding workflow process model is made up of 16 activities, such as START, END, collect reviews, decide, invite additional reviewer, invite reviewers, get review 1, get review 2, get review 3, get review X, time- out 1, time-out 2, time-out 3, time-out X, accept, and reject.

Table 1. Next Workflow-Activities from Time-Out 2

Input Activity	Next Activity	Probability
time-out 2	time-out 1	24%
time-out 2	get review 1	14%
time-out 2	get review 3	12%
time-out 2	time-out 3	11%
time-out 2	invite additional reviewer	7%
time-out 2	invite reviewers	4%
time-out 2	reject	4%
time-out 2	collect reviews	3%
time-out 2	time-out 2	3%
time-out 2	accept	3%
time-out 2	END	2%
time-out 2	get review X	2%
time-out 2	get review 2	2%
time-out 2	START	1%
time-out 2	time-out X	1%

Using the temporal workcases preprocessed from the original event log dataset, we first applied the deep learning neural network model to them. We designed and implemented a simple prototype by the Google's Tensorflow machine learning framework and its Python programming. For the neural network model, we set two hidden LSTM layers and each layer has 128 nodes with 50 of batch size and 200-epoch of training repetition.

For the sake of understandable predictions, we firstly set the probability threshold to 1%. We predict the next workflow-activities of each workflow-activity in the Review Example Large" workflow process model through the following approach that is to predict the next workflow-activities without former sequences. In this approach, the deep learning neural network model only gets one input workflow-activity and predicts the next workflow-activities of the input. Thus, through this approach, we were able to see the overall relationships of the input workflow-activity to others.

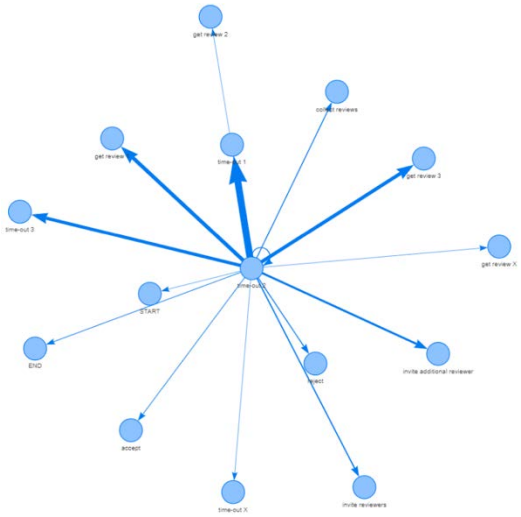


Fig. 2. Visualizing Next Workflow-Activities with their Probabilities from Time-Out 2

Table 1 and **Fig. 2** are listed and visualized those next workflow-activities with their probabilities from the specific Time-Out 2 workflow-activity, respectively, as an example. As you can see, the prediction results of the Time-Out 2 workflow-activity show 15 candidates as the next workflow-activities. From this situation, we intuitively guess that the Time-Out 2 workflow-activity ought to be associated with many different control-paths in the corresponding workflow process model. In **Fig. 2**, the centered node is the input-featured workflow-activity; the satellited nodes represent the output-featured workflow-activities, and the thickness -level of the directed edges between the centered and the satellited are illustrating the numerical value of probabilities.

5. Conclusions

In this paper, we designed and implemented a LSTM-based deep learning neural network model that is able to predict the next workflow-

activities from each of the associated workflow-activities in the workflow process model. To fulfill the series of predictions, firstly we preprocessed the original dataset of event logs and prepared the temporal workcases dataset to be used for training the proposed LSTM-based deep learning neural network model. After training the model, we applied the model to the prediction phase of the approach. Finally, we carried out an operational experiment on the synthetic dataset of the Review Example Large workflow process instances' execution event histories and verified the correctness of the deep learning approach proposed in this paper. As the future works of this paper, we will try to find an engineering approach to apply the deep learning approach and its experimental results to the workflow process remodeling and reengineering works.

References

[1] Hyun Ahn, Kwanghoon Pio Kim, "Formal approach for discovering work transference networks from workflow logs", *Information Sciences*, Vol. 515, pp. 1-25, 2020. DOI: 10.1016/j.ins.2019.11.036

[2] Kyoung-Sook Kim, et al., "A Conceptual Approach for Discovering Proportions of Disjunctive Routing Patterns in a Business Process Model", *KSII Transactions on Internet and Information Systems*, Vol. 11, No. 2, pp. 1148-1161, 2017. DOI: 10.3837/tiis.2017.02.030

[3] Dinh-Lam Pham, Hyun Ahn, Kwanghoon Pio Kim, "Discovering Redo-Activities and Performers' Involvements from XES-Formatted Workflow Process Enactment Event Logs", *KSII Transactions on Internet and Information Systems*, Vol. 13, No. 8, pp. 4108-4122, 2019. DOI: 10.3837/tiis.2019.08.016

[4] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[5] J. Evermann, J.-R. Rehse, and P. Fettke, "A deep learning approach for predicting process behaviour at runtime," in International Conference on Business Process Management. Springer, 2016, pp. 327–338.

[6] N. Tax, I. Verenich, M. La Rosa, and M. Dumas, "Predictive business process monitoring with lstm neural networks," in International Confer- ence on Advanced Information Systems Engineering. Springer, 2017, pp. 477–492.

[7] N. Navarin, B. Vincenzi, M. Polato, and A. Sperduti, "Lstm networks for data-aware remaining time prediction of business process instances," in 2017 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE, 2017, pp. 1–7.

[8] Seong-Hun, Hyun Ahn, Kwanghoon Pio Kim, "LSTM-based Business Process Remaining Time Prediction Model Featured in Activity-centric Normalization Techniques," JICS, Vol. 21, No. 3, pp. 83-92, 2020. DOI: 10.7472/jksii.2020.21.3.83

[9] M.-J.Park and K.-H.Kim, "Xwell: XML-based workflow event logging mechanism and language for workflow mining systems," in International Conference on Computational Science and its Applications, Springer, 2007, pp. 900–909.

[10] M. zur Muehlen and K. D. Swenson, "BPAF: a standard for the interchange of process analytics data," in International Conference on Business Process Management, Springer, 2010, pp. 170–181.

[11] X. W. Group et al., "IEEE standard for extensible event stream (XES) for achieving interoperability in event logs and event streams," IEEE Std, vol. 1849, 2016.

KSII The 12th International Conference on Internet (ICONI) 2020.

Sign language translation performance using Transformer according to hyperparameter

San Kim, Yoonyoung Jeong, Han-Mu Park, Hyedong Jung
Artificial Intelligence Research Center, Korea Electronics Technology Institute
Seongnam, South Korea
[e-mail: {kimsan0622, yoonyj, hanmu, hudson}@keti.re.kr]
*Corresponding author: San Kim

Abstract

In this paper, we experimented on how the performance changes according to hyper parameters in Korean sign language translation using Transformer. In particular, among the hyperparameters, the performance according to the number of layers and number of attention heads was measured. As a result of the measurement, the higher the number of attention heads, the higher the performance. It was confirmed that the performance increased as the number of layers increased, but the performance decreased when the number of layers was greater than a certain value.

***Keywords*:** Sign language translation, Transformers, Human keypoints, hyperparameter

1. Introduction

Various models and methods[1] have been developed for sign language translation over the past few years. In particular, the use of Transformers[2] and human keypoints has greatly improved the performance of sign language translation. However, they did not go into detail about the translation performance of the sign language according to Transformer's hyperparameters.

Therefore, in this paper, we measure and analyze the performance according to the number of layers and the number of attention heads among many hyperparameters of the Transformer.

2. Sign language Transformer

2.1 Sign language Transformer based on position of human keypoints

The sign language translation task is a task that converts a sign language video into a corresponding sentence. Transformer consists of an encoder and a decoder. Body, hands, and face keypoints are extracted from the video and the extracted values are used as input to the encoder. The encoder extracts the contextual embedding of the input feature vector sequence via self-attention. The sentence corresponding to the video is tokenized and input to the decoder. The decoder extracts the contextual vectors of the input sequence via self-attention and expresses the decoder inputs as the outputs of the encoder through cross-attention.

2.2 Position Normalization

The positional variation of the extracted human keypoints of the frames is large enough to affect the performance. The reason is that each person has different physical characteristics such as height and arm length, and the distance from the camera is different. Therefore, in order to

minimize this variation, various normalization methods are used in sign language translation using human keypoints. We used the length of the neck to normalize the position of all human joints.

3. Experiments

We set all hyperparameters the same except for the number of layers and the number of attention heads. The results of all experiments are the average of the five trials each. **Table 1-3** show BLEU4, METEOR, and RougeL scores, respectively.

For BLEU4, if there is only one layer, the higher the number of attention heads, the higher the performance. However, when the number of layers was 2, there was little difference in performance according to the number of attention heads, and when the number of layers was 4, the performance decreased as the number of attention heads increased.

Table 1. BLEU4 Score

		Number of layers		
		1	2	4
Number of attention heads	1	0.7464	0.7678	0.7614
	2	0.7582	0.7739	0.7631
	4	0.7654	0.7425	0.7480
	8	0.7680	0.7784	0.6420

Table 2. METEOR Score

		Number of layers		
		1	2	4
Number of attention heads	1	0.7932	0.8141	0.7982
	2	0.7957	0.8164	0.8032
	4	0.7965	0.7844	0.7893
	8	0.8108	0.8182	0.8173

Table 3. RougeL Score

		Number of layers		
		1	2	4
Number of attention heads	1	0.8086	0.8296	0.8132
	2	0.8098	0.8315	0.8188
	4	0.8123	0.8006	0.8048
	8	0.8273	0.8242	0.8326

The METEOR score showed the same trend as BLEU4 when the number of layers was 1 or 2. When the number of layers was 3, unlike BLEU4, there was no difference in performance according to the change in the number of attention heads. In the case of the RougeL score, the higher the number of attention heads, regardless of the number of layers, the higher the performance.

When the number of layers is 4 and the number of attention heads is 8, the RougeL score is the highest, but the BLEU4 and METEOR scores are not. When all the performance metrics were considered, the highest performance was obtained when the number of layers was 2 and the number of attention heads was 8.

4. Conclusions

In this paper, we measured the performance according to the number of layers and number of attention heads in the sign language translation task. As a result of the experiment, in general, the higher the number of attention heads, the higher the performance. However, when the number of layers exceeded the appropriate size for the task and dataset, the performance deteriorated.

Acknowledgments

This work was supported by Institute for Information & Communications Technology Promotion (IITP) and National IT Industry Promotion Agency (NIPA) grant funded by the Korea government (MSIT) 2017-0-00255 and S1601-20-1034)

References

[1] S. Ko, C. Kim, H. Jung and C. Cho, “Neural sign language translation based on human keypoint estimation,” *Appl. Sci.,* Vol 9, No. 13, 2019.

[2] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin, “Attention is all you need,” *Advences in Neural Information Processing Systems* 30, 2017.

[3] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh, “Openpose: Realtime multi-person 2d pose estimation using part affinity fields,” in *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 2019.

KSII The 12th International Conference on Internet (ICONI) 2020.

Graph Transformer for Motion Prediction

San Kim, Yoonyoung Jeong, Saim Shin, Hyedong Jung
Artificial Intelligence Research Center, Korea Electronics Technology Institute
Seongnam, South Korea
[e-mail: {kimsan0622, yoonyj, sishin, hudson}@keti.re.kr]
*Corresponding author: San Kim

Abstract

In this paper, we introduce graph transformers that combine graph convolution using features of nodes and adjacency matrix and attention calculation in the time domain of transformers. The proposed graph transformers simultaneously perform an operation in the spatial domain and an operation in the time domain, and unlike existing spatial transformers, connection information between nodes can be explicitly input. Explicitly entering the connection information between nodes shows good results in motion prediction where the connection information of joints is important. Graph transformers showed low euler error in repetitive simple motion prediction tasks such as walking and jumping.

***Keywords*:** Graph convolution, Motion prediction, Transformers, Spatial Transformers, Graph Transformers

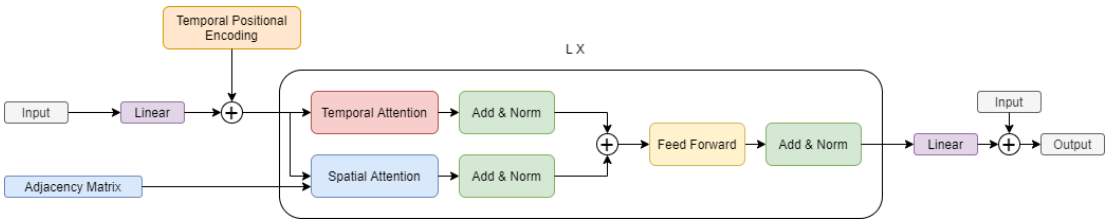


Fig. 1. Graph Transformers

1. Introduction

Various models and methods have been developed for 3D motion prediction over the past few years. There are studies using graph convolution networks and GRU [1], and studies that transform transformers [2].

In particular, the research using transformers developed that the existing transformers calculate attetion only in the time domain, and thus, the study was conducted to simultaneously calculate attention in the spatial and temporal domains. We propose graph transformers that explicitly input connection information between nodes in the spatial domain and use this information to calculate spatial attention.

2. Graph Transformers

The proposed graph transformer is an explicit graph convolution operation added to the spatial attention calculation in spatio-temporal transformers [2] that added spatial attention in existing transformers.

Fig. 1 shows the structure of Graph Transformers. Graph Transformers uses an

adjacency matrix to calculate the spatial attention of existing spatio-temporal transformers. After calculating the spatial score by using the input, the spatial attention is calculated by applying softmax to the spatial score. The final spatial representation is obtained by adding the adjacency matrix to the spatial attention and multiplying it by the value.

Through this process, the connection between nodes is explicitly used for spatial attention calculation, and if an input tensor is not input in the process of calculating spatial attention, it is an operation like graph convolution. It is the same as existing spatio-temporal Transformers and Transformers except for the process of calculating spatial attention.

3. Experiments

CMU motion capture data was used as the 3D motion prediction dataset, and downsampled to 30 frames per second. We experimented using categories of basketball forward dribble, walking, jumping, and soccer, and we used 5 motions per category for the train data set and 1 motion per category for the test data set. In the test process, 2 seconds, that is, 60 frames was used as a seed frame. As the angle representation, the product of the axis and angle in the axis-angle representation was used, and the axis-angle in each joint frame was used.

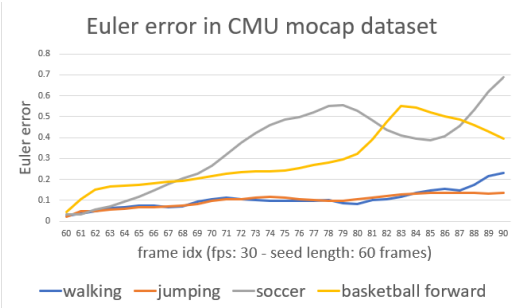


Fig. 2. Euler error in 3D motion prediction

Fig. 2 shows the euler error between the ground truth motions and the predicted motion over time. The euler error was measured by converting the rotation matrix in the joint frame into a global frame, because the error of the parent node is propagated to the children node. When drawing the position of the joint in the actual 3D space, it is reasonable to calculate the euler error in the global frame because the position in the global frame is drawn.

In the case of relatively regular walking and jumping, errors do not accumulate well over time. However, for relatively complex motions such as soccer and basketball forward dribble, errors can be seen to accumulate rapidly.

4. Conclusions

In this paper, we proposed Graph Transformers and introduced a 3D motion prediction method using Graph Transformers. Graph Transformers can explicitly use the connection information between nodes in the attention calculation process.

We plan to make 3D motion prediction with an encoder-decoder structure using Graph Transformers. In addition, we plan to conduct research to generate 3D motion using natural language.

Acknowledgments

This work was supported by Institute for Information & Communications Technology Promotion (IITP) and National IT Industry Promotion Agency (NIPA) grant funded by the Korea government (MSIT) 2017-0-00255 and S1601-20-1034)

The data used in this project was obtained from mocap.cs.cmu.edu.

References

[1] Maosen Li, Siheng Chen, Yangheng Zhao, Ya Zhang, Yanfeng Wang, Qi Tian, “Dynamic Multiscale Graph Neural Networks for 3D Skeleton Based Human Motion Prediction,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2020, pp. 214-223, 2020.*

[2] Emre Aksan, Peng Cao, Manuel Kaufmann, Otmar Hilliges, “A Spatio-temporal Transformer for 3D Human Motion Prediction,” *ArXiv preprinted ArXiv:2004.08692, 2020.*

Multimodal Emotion Recognition on a Korean Multimodal Video Dataset Using K-means Clustering

Hyeontae Seo, Jin Yea Jang, Saim Shin, Hyedong Jung and Minyoung Jung*
Artificial Intelligence Research Center, Korea Electronics Technology Institute
Seongnam, South Korea
[e-mail: {dchs504, jinyea.jang, sishin, hudson, minyoung.jung}@keti.re.kr]
*Corresponding author: Minyoung Jung

Abstract

Emotion recognition has gained significant attention in various applications, and many researchers have conducted emotion recognition using a variety of input modalities. Although recently researchers have studied multimodal emotion recognition, emotion recognition is still challenging because even humans are often confused about others' emotions. We introduce an unsupervised learning approach of multimodal emotion recognition with k-means clustering because even human-labeled emotion data might be inaccurate. We evaluate our approach on audio-textual data from the Korean multimodal video dataset.

***Keywords*:** multimodal emotion recognition, k-means clustering, deep learning, unsupervised learning

1. Introduction

Emotion recognition is utilized in many fields, including human-computer interaction. Although many researchers have performed emotion recognition on an input modality such as facial expressions, speech or written text, detecting emotions is still one of challenging tasks because humans use a lot of indirect and direct means to represent their emotions.

To overcome the confines of emotion recognition on a single modality, recently many research works have been focused on multimodal emotion recognition because it has richer information. Although multiple input modalities are employed in emotion recognition, it is still a demanding task to detect accurate human emotions because even humans are often confused about others' emotions. It means even human-labeled data in emotion recognition datasets might be inaccurate.

Unsupervised learning is a machine learning approach mainly to deal with unlabeled data. We adopt k-means clustering [1], which is one of the most popular unsupervised learning techniques, to mitigate the problem of emotion recognition data. We conduct experiments of multimodal emotion recognition on audio and text data from the Korean multimodal video dataset [2] using k-means clustering.

2. Dataset

The total 78,215 audio-textual data in the Korean multimodal video dataset are extracted, and their target emotions are angry, sad, neutral, surprise, dislike, fear, contempt and happy. The data distribution by emotions is shown in **Table 1**.

This work was partly supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (2017-0-00255, Autonomous Intelligent Digital Companion Framework and Application) and National IT Industry Promotion Agency (NIPA) grant funded by the Korea government(MSIT) (S1601-20-1034, Development of artificial intelligence based empathy service technology for emotional stability).

Table 1. Data distribution by emotions

Emotion	Number of data
Angry	7,875
Sad	7,969
Neutral	11,455
Surprise	12,065
Dislike	17,765
Fear	5,079
Contempt	2,526
Happy	13,481

3. Multimodal Emotion Clustering

3.1 Audio Feature Extraction

We convert the audio-visual data (mp4 files) into the audio data (wav files), and extract eight audio features (the mean and standard deviation, the mean and standard deviation by calculating the root-mean-square for each frame, the number of counted parts when silent, the maximum and standard deviation of autocorrelation on center-clipped frames, and the high-pitch).

3.2 Text Feature Extraction

We extract text features through a sequence-to-sequence (seq2seq) autoencoder which is trained to copy its input to its output. An autoencoder has an encoder to learn a reduced representation from input, and a decoder to generate a representation as close as possible to its original input from the reduced representation. For our experiments, we utilize the hidden features produced by the encoder as the text features.

3.3 K-means Clustering

For unsupervised learning of multimodal emotion recognition, we employ a well-known clustering method, k-means. K-means clustering is a distance-based algorithm which identifies k centroids, and allocates all data points to the nearest cluster while keeping the centroids as small as possible. We combine different features extracted from two input modalities, speech and texts, and then evaluate multimodal emotion recognition on the combined features using k-means clustering.

4. Experimental Results

Fig. 1 shows the feature vectors which are projected onto a two-dimensional space via Principal Component Analysis (PCA). After our multimodal clustering, the predicted emotions of the Korean multimodal video data look clustered, unlike the labeled emotions.

We sampled some utterances whose labeled and clustered emotions are different, and the clustered emotion looks more accurate than the labeled one when we checked the audio and text. **Table 2** shows the utterance samples with its labeled and clustered emotions.

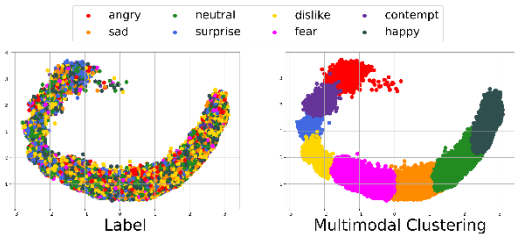


Fig. 1. PCA of labeled and clustered emotions

Table 2. Labeled and clustered emotions of samples

Utterance	Emotion	
	Labeled	Clustered
"네..." (“Yes…”)	Neutral	Sad
"바이러스 아니야?" (“Isn’t it a virus?”)	Neutral	Fear
"하고 싶은대로 해." (“Do whatever you want.”)	Sad	Neutral

5. Conclusions

Emotion recognition is one of challenging and important tasks in many fields. In this paper, we perform multimodal emotion recognition on audio and text data from the Korean multimodal video dataset using k-means clustering for unsupervised learning. Our experiments show that the predicted emotions of the Korean multimodal video data look clustered, while the labeled emotions do not.

References

[1] James B. MacQueen, “Some methods for classification and analysis of multivariate observations,” in *Proc. of the Fifth Symposium on Math, Statistics, and Probability*, 1967.

[2] Korean multimodal video dataset, “https://aihub.or.kr/aidata/137,” retrieved Sep. 2020.

Environmental Multi-Sensor Module for Flexible Combination

Myeongseop Kim, Tai-Gil kwon, Seho Park and Kyung-Taek Lee
Korea Electronics Technology Institute
Seoul, South Korea
[e-mail: myeongseopkim@keti.re.kr, tgkwon@keti.re.kr, sehopark@keti.re.kr, ktechlee@keti.re.kr]
*Corresponding author: Kyung-Taek Lee

Abstract

The development of machine learning and deep learning technology has enabled environmental data to find meaningful signs like condition prediction and detection. Therefore, the need to collect various data from general-purpose sensors rather than special-purpose sensors has increased. This paper implemented a novel environmental multi-sensor module for flexible combinations related to general-purpose sensors. The proposed system increased flexibility by combining data from six general-purpose sensors (illumination, Air quality, temperature, humidity, CO2, fine dust, Motion, IR) and communicates data using I2C, UART, and Serial interfaces. Also, an experimental configuration that stores 9,600 databases per day has been established which can be used for anomaly prediction and detection models.

***Keywords*:** Environmental Multi-Sensor, Embedded Artificial Intelligence, Anomaly Detection, Flexible Combination, Lightweight Deep Learning

1. Introduction

More and more environmental sensors are used to predict and detect symptoms in factories, houses, and public spaces. Also, as machine learning and deep learning has been gradually developed, collecting multiple data becomes more important for creating a novel prediction model. Tengyue and Fong [1] presented a model to predict the number of passengers on a public bus by applying a neural network to carbon dioxide data. Uday et al. [2] suggested an intelligent human counting system using gas data through environmental sensing in closed indoors by adapting the CNN model. As these methods collected data by special-purpose sensors of specific space, It is hard to apply their sensing data to different places. Therefore, this paper implements an environmental multi-sensor module that can collect various environmental data in different spaces such as houses, schools, hospitals, factories by utilizing general-purpose sensors (e.g., Illumination, Air quality, Temperature, Humidity, CO2, Fine dust, Motion, IR).

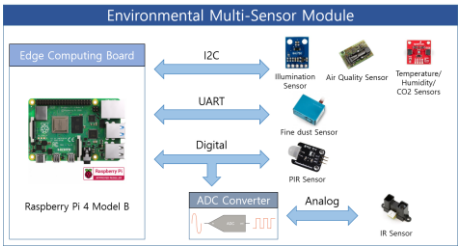


Fig. 1. Illustration of environmental multi-sensor module for flexible combination

As the environment multi-sensor module requires various environment sensing data and

This work was supported by the IT R&D program of MOTIE/KEIT. [20010525,Development of intelligent complex environmental sensor technology capable of flexible reconfiguration]

high computational cost, the edge computing paradigm is applied to implement state monitoring and abnormal state detection very efficiently [3]. Gang et al. [4] applied the edge computing framework to real-time fault diagnosis and dynamic control research of the rotating machine using the sensor to reduce data storage and computing resources. In this study, an environmental sensor module was created based on an edge computing board to process and compute various environmental data.

2. Environmental Multi-Sensor Module for Flexible Combination

2.1 Edge Computing Board

Baotong et al. [5] proposed that Raspberry Pi can be used as a data fusion and edge computing node. Therefore, in this study, Raspberry Pi consisting of Cortex-A72 64bit CPU was used as an edge board.

2.2 Environmental Multi-Sensor Module

The designed module has combined six kinds of general-purpose sensors (Illumination; Air-Quality; Temperature/Humidity/CO2; Fine dust; Motion; Infrared light) on the edge board (**Fig. 1, 2**) to collect synchronized sensing data by connecting various sensors fluidly. The fabricated sensor modules can support all I2C, UART, and Serial(Digital, Analog) communication interfaces.

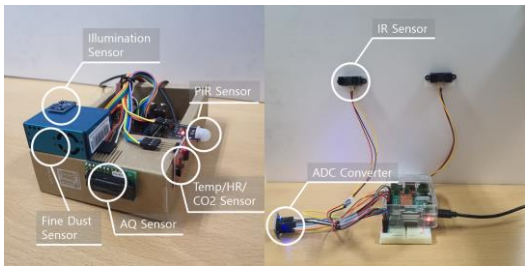


Fig. 2. Environmental multi-sensor module

2.3 Collected Datasets

The proposed environment multi-sensor module for flexible combinations collected environment data for indoor space. As shown in **Table 1**, The database consists of date, time, and seven types of environmental a data, and the number of workers in the experiment space. The environmental multi-sensor module collected synchronized sensor data in approximately every 10 seconds. In other words, a database of 9,600 data can be collected in one day. The collected data were stored on the server in real-time by the network. This can be used as learning data for a prediction and detection model for future work.

Table 1. Environmental data parameters

Data	Range
Date	2020.xx.xx
Time	00:00:00 ~ 23:59:59
Illumination	1 ~ 65535 lx
eCO2	400 ~ 60000 ppm
TVOC	0 ~ 60000 ppm
CO2	0 ~ 10000 ppm
Temperature	-40 ~ 125 dgree
Humadity	0 ~ 100 %RH
Paticle	Different by particle size
Number of worker	Number

3. Conclusions

In this paper, we designed a novel environmental multi-sensor module for a flexible combination of adapting general-purpose sensors. The collected datasets by the proposed module can be used to predict and verify environmental symptoms, such as human counting or indoor pollution, in the future.

References

[1] Li, Tengyue, and Simon Fong. “Counting Passengers in Public Buses by Sensing Carbon Dioxide Concentration: System Design and Implementation.” *In Proceedings of the 2018 2nd International Conference on Big Data and Internet of Things*, 2018.

[2] Kamal, Uday, et al. “Intelligent Human Counting through Environmental Sensing in Closed Indoor Settings.” *Mobile Networks and Applications*, pp. 1-17, 2019.

[3] Shi, Weisong, and Schahram Dustdar. “The promise of edge computing.” *Computer*, Vol 49, No. 5, pp. 78-81, 2016.

[4] Qian, Gang, et al. "Edge computing: A promising framework for real-time fault diagnosis and dynamic control of rotating machines using multi-sensor data." *IEEE Sensors Journal*, Val 19, No. 11 pp. 4211-4220, 2019.

[5] Chen, Baotong, et al. “Edge Computing in IoT-Based Manufacturing” *IEEE Communications Magazine*, Val 56, No. 9, pp. 103-109, 2018.

Tactile image conversion technique of smart media device for visually impaired

Youngmin Kim, Jisoo Hong, Sunghee Hong, and Jinsoo Jeong
Hologram research center, Korea Electronics Technology Institute
Seoul, South Korea
[e-mail: rainmaker@keti.re.kr, jhong@keti.re.kr, shhong@keti.re.kr, j.jeong@keti.re.kr]
*Corresponding author: Youngmin Kim

Abstract

This paper describes tactile image conversion technique of smart media device for visually impaired. This smart media device consists of multiple braille cells which are powered by actuator technology. Each cell operates with fast refresh rate, so lots of contents could be expressed even image files with low resolution. We proposed image conversion techniques which could be useful for this smart device, the main purpose of this was to convert 3D model of articles in the museum from 3D scan techniques into tactile images. We hope this method could be applied into many applications such as navigation or image books for visually impaired.

***Keywords*:** Visually impaired person, barrier-free device, indoor navigation, text to braille engine, tactile contents authoring tool

1. Introduction

Accessing museums or art galleys has been difficult for people who are blind or visually impaired people. Nowadays, there are approximately 1.3 billion people living with blindness or visually imparied globally, so visiting to a museum or art galley has the potential to make them feel another experiences. One of the means of assistance in the museum, audio guides and audio descriptions, could be one candidate for them, however, more innovative approaches are needed for visually impaired person [1, 2]. One of the approaches is to use multi-modal properties such as a tactile device or exhibition with sound and smell, however much researches still needed. With the help of relatively high-speed actuators, several companies have focused on researching tactile devices. In this paper, we proposed image conversion techniques which could be useful for such tactile devices.

2. Proposed method

2.1 Work flow of the proposed method

For expressing tactile image, the original objects (e.g. articles in museum) were converted into high-quality 3D modeling scan data. However, until now, the degree of freedom to express the articles was limited with tactile device. For example, the up-to-date tactile device from DOT incorporation [3] consists of hundreds or thousands of actuators (each actuator consisted of 8 dots.); total resolution of tactile device was limited in 10K. The level of these resolutions is relatively low compared with commercially available two-dimensional display, so the tactile image conversion technique will be needed. So, we designed contents authoring

This research is supported by Ministry of Culture, Sports and Tourism and Korea Creative Content Agency (Project Number: R2020040214, “Development of multi-modal exhibition guide terminals and authoring technology for smart viewing of visually impaired people”).

software that could convert into simple dot images from sophisticated original objects. **Fig. 1** shows the schematic of the proposed method. Once the objects were captured by high performance 3D scanner, the 3D model data was converted into simple two-dimensional dot-images with high frequency component. The dot image could be expressed with pre-determined resolution according to the tactile device resolution, so each tactile device could provide two-dimensional dot image to the visually impaired. The conversion process was performed by edge detection techniques.

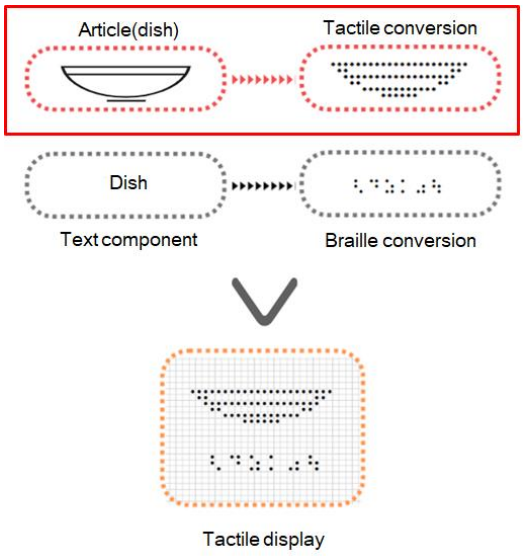


Fig. 1. Work flow of the proposed method

2.2 Manipulation of the conversion image

However, the errors could be occurring in the conversion process. Furthermore, the total resolution of the conversion image is under 10K, so edge-oriented image contents from conversion process could not be included with original intention. Therefore, we added the function that could manipulate each dot pixel from conversion process. This process is similar with the function of alpha channel in Photoshop, so users such as curators in museum were able to revise the conversion image. **Fig. 2** shows the example of the manipulation process of the conversion image.

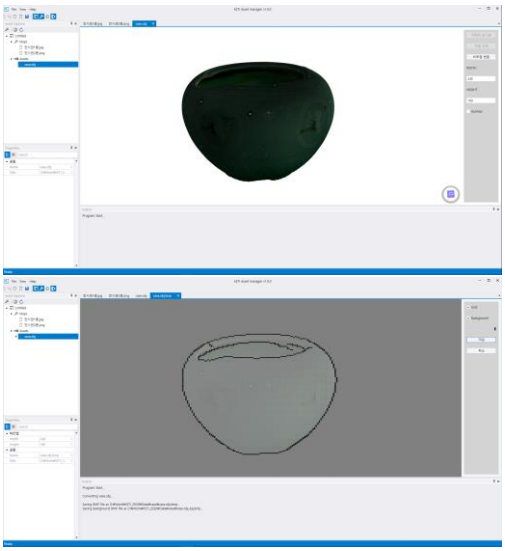


Fig. 2. An example of the manipulation process of the conversion image

3. Conclusions

In this paper, we proposed an image conversion technique of tactile smart media for visually impaired. The articles could be inverted into 3D computer graphics model, and it could be converted into two-dimensional edge-oriented dot-images. The images could be manipulated by curators, so the image conversion process is more adaptive. We hope this method could be applicable into the museum or art gallery for visually impaired person.

References

[1] S. Asakawa, J. Guerreiro, D. Ahmetovic, and K. Makoto Kitani, "The present and future of museum accessibility for people with visual impairments," in *the 20th International ACM SIGACCESS Conference*, pp. 382-384, 2018.

[2] D. Leithinger, S. Follmer, A. Olwal, and H. Ishii, "Physical telepresence: shape capture and display for embodied, computer-mediated remote collaboration," Proceedings of the 27th annual ACM symposium on User interface software and technology, pp. 461-470, 2014.

[3] Dotincorporation, http://dotincorp.com.

KSII The 12th International Conference on Internet (ICONI) 2020.

Secure IoT Smart-sensors based on Energy Harvesting in the Power Plant

Jinhee Park and Yongho Kim
Korea Electronics Technology Institute
Seongnam, South Korea
[e-mail: pjhe@keti.re.kr, yhkim@keti.re.kr]
*Corresponding author: Jinhee Park

Abstract

With the proliferation of industrial fields applied IoT technologies such as smart factories, smart buildings and energy plants, the demand for energy harvesting-based smart sensors that facilitate power supply and system maintenance has been increased. Since smart sensors can be used in these industrial domains for a wide area of several kilometers, wireless communication that satisfies the security level required by industrial sites is demanded. In this paper, we designed a method for secure communication between IoT smart sensors in the power plants.

***Keywords*:** IoT, secure communication, smart sensor, power plant

1. Introduction

In the IoT environment, various devices ranging from 8-bit low-end devices to 64-bit high-spec devices are interconnected to perform communication. However, due to the openness of IoT devices, security threats such as eavesdropping on wireless communication and privacy invasion are increasing. Due to these characteristics, security threats in the IoT environment can cause more serious problems [1]. In specially, smart sensors can be used in these industrial domains for a wide area of several kilometers, wireless communication that satisfies the security level required by industrial sites is demanded. In this paper, we consider security requirements of smart sensors network and propose secure structure of IoT smart-sensors and DCU installed in the power plant by message authentication and integrity mechanism.

2. Security Requirements of IoT Smart-sensors

In the IoT Smart-sensor network, the guaranteeing the confidentiality and integrity of messages is inevitably considered the most important security factor in the IoT communication environment. ZigBee, an IoT communication protocol, uses AES-CCM* to perform message integrity and sender authentication. Message Authentication Code is an encryption technology that provides authentication for this as well as message integrity. By sending a message and an authentication tag together, it verifies that the content of the message has not been unjustly changed during transmission and is safely received.

This work was supported by the Energy Efficiency & Resources Core Program of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) granted financial resource from the Ministry of Trade, Industry & Energy, Republic of Korea (20192010106790)

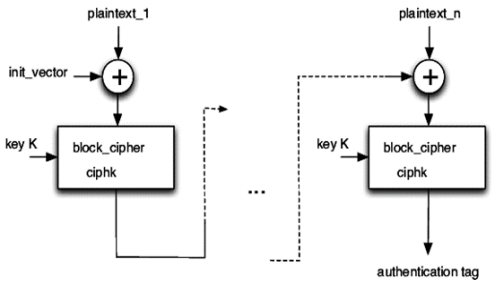


Fig. 1. CBC-MAC block diagram

Counter with CBC-MAC (CCM) is a generic authenticated encryption block cipher mode. CCM is defined for use with 128-bit block ciphers, such as the Advanced Encryption Standard (AES) [2]. CBC-MAC is the most well-known algorithm and generates an authentication tag using the CBC mode.

3. Secure IoT Smart-sensors

As the name suggests, CCM mode combines the well known CBC-MAC with the well known counter mode of encryption. These two primitives are applied in an "authenticate-then-encrypt" manner, that is, CBC-MAC is first computed on the message to obtain a tag t; the message and the tag are then encrypted using counter mode. One key insight is that the same encryption key can be used for both, provided that the counter values used in the encryption do not collide with the (pre-)initialization vector used in the authentication.

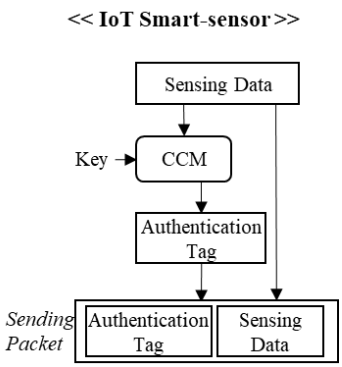


Fig. 2. Secure IoT Smart-sensors structure

A MAC algorithm is a cryptographic algorithm that computes a complex function of a plaintext and a secret key; the resulting MAC value is typically appended to the plaintext to protect its authenticity. CBC-MAC is a MAC algorithm based on a blockcipher; it is derived from the Cipher Block Chaining (CBC) mode of operation, which is a mode for encryption.

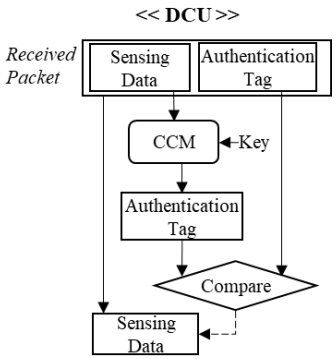


Fig. 3. Secure DCU structure

4. Conclusions

In this paper, we propose a method for secure communication between IoT smart sensors in the power plants by CCM. Since wireless communication that satisfies the security level required by industrial sites is demanded. It is thought that the IoT smart sensor with security applied will be able to contribute to the secure communication within the power plant.

References

[1] Minwoo Kim and Taekyoung Kwon, “Analysis of Research Trend and Performance Comparison on Message Authentication Code”, Journal of KIISE, Vol. 43, No. 11, pp. 1245-1258, 2016.

[2] D. Whiting, et all, “Counter with CBC-MAC(CCM)”, IETF RFC 3610, 2003.

[3] Seonghyeon Gong and Chang-hoon Lee, “Analysis of Authenticated Encryption”, Communications of the Korean Institute of Information Scientists and Engineers, Vol 35, No. 1, pp. 23-31, 2017.

[4] Taehwan Park, Hwajeong Seo, Shinwook Heo, and Howon Kim, “The Recent Research Trend and Prospective on Authenticated Encryption Schemes”, KIISE Transactions on Computing Practices, Vol 24, No. 10, pp. 563-568, 2018.

Detection of wheezing on lung sound using deep learning

Saetbyeol Yu and Seho Park
Korea Electronics Technology Institute
Seoul, South Korea
[e-mail: sbyu@keti.re.kr, sehopark@keti.re.kr]

Abstract

This paper describes how to detect wheezing on lung sound. Wheezing is a common sound in various lung diseases. If wheezing is properly detected during auscultation, it is possible to prevent various lung diseases. A smartphone can record lung sound using an external device based on a microphone. The recorded data is transmitted to the lung sound analyzing server, and the server analyzes recorded data using a pre-trained model. The server pre-processes the recorded data and analyzes it. After pre-processing the recorded data, the server can classify which include wheezing or not. To train the model, we use the Convolution Neural Network(CNN) and the accuracy of the model is 87% on the test data set.

***Keywords*:** Lung Sound, Wheeze, Respiration cycle, Deep learning, AI

1. Introduction

As the demand for remote medical treatment increases, services are being developed to measure and collect a user's physical condition using digital devices, and to measure an individual's health condition using the collected data. In particular, in the case of lung disease, since the lung condition can be easily diagnosed through auscultation, a self-examination method is needed that allows the general public to recognize and judge the possibility of lung disease early. In this paper, we design and implement a deep learning-based lung sound analysis system that can detect wheezing from audio-recorded lung sounds. The developed model determines whether Wheezing is included in the recorded lung sound using a microphone-based device that can be connected to a smartphone, and the user can easily check the health of the lungs at home.

2. Data Preprocessing

2.1 Wheezing sound

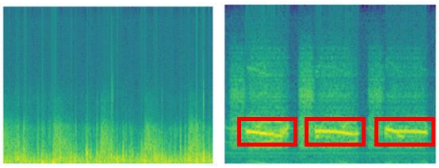


Fig. 1. Wheezing features in Spectrogram

Wheezing is the sound you heard as air passes through narrowed bronchi when you breathe. Wheezing is a sound that can be heard in most lung-related diseases, if you can hear the wheezing during auscultation, lung disease can be suspected. Comparing the spectrogram of normal lung sound and abnormal lung sound, as shown in **Fig. 1**, where wheezing occurs, a horizontal line is displayed in the spectrogram. In

This research was supported by the Ministry of Science and ICT(MSIT) and Institute for Information & communications Technology Promotion (IITP). [2019-0-00711, Development of Artificial Intelligence Chest Sound Analysis Technology and Chest Sound Data Operation System Technology].

this paper, we design and implement a model that detects wheezing using spectrogram images.

2.2 Preprocessing

As described in 2.1, pre-processing is required to detect wheezing in the spectrogram. Wheezing may occur in all respiratory cycles or irregularly depending on the patient. Therefore, the input unit of the training model was defined as one cycle of inspiration and exhalation. The lung sounds recorded with a digital stethoscope contain not only lung sounds, but also noises such as heart sounds and background sounds. Since the lung sound signal is mainly distributed in the low frequency band, a bandpass filter of 60Hz ~ 600Hz was used to remove noise. The signal passed through the filter was enveloped and the respiratory cycle was calculated by finding the minimum and maximum points. The average respiratory cycle of the datasets is 3.7 seconds. To fix the size of the input, create a spectrogram of one breath cycle in a 5-second window. If the signal is less than 5 seconds, 0 is added as padding, and if the signal is greater than 5 seconds, the remaining signal is truncated.

3. Datasets and CNN Model

We used lung sound datasets provided in [1] as training data. There are some noisy data in the data set provided. We excluded noisy data from the training dataset to build the model with high accuracy. The providing dataset is labeled for wheezing and crackle. After segmenting the original audio data into a respiration cycle, converting it into a spectrogram, selecting only the data labeled as wheezing and normal were used as training data. Also, using the developed device, the normal lung sound of 20 seconds 216 were collected and added to the training dataset [2]. The number of training datasets labeled by dividing by respiration cycle is 3720 normal and 807 abnormal.

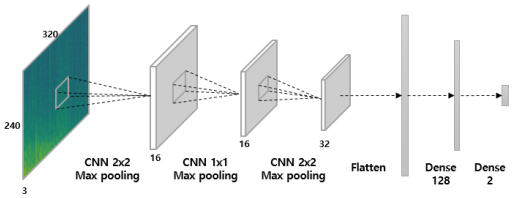


Fig. 2. Structure of CNN Model

A CNN model was designed to detect wheezing in the spectrogram image. The designed CNN model is shown in **Fig. 2**[3].

The training result of the designed model is shown in **Fig. 3**. The accuracy of the test set is 87%.

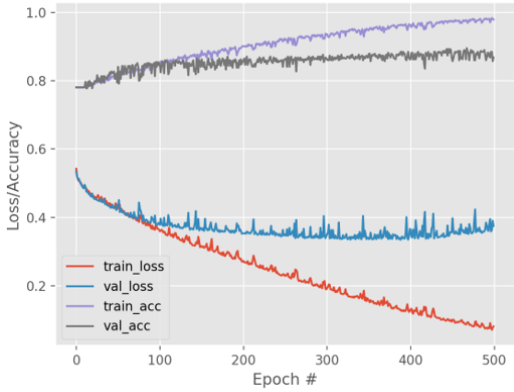


Fig. 3. Accuracy and Loss of Model

4. Conclusions

In this paper, we designed and implemented a deep learning model that detects wheezing from recorded lung sound. The accuracy of the implemented model is 87%. In the future, we plan to conduct research that can be utilized in remote medical services using smart devices.

References

[1] https://www.kaggle.com/vbookshelf/respiratory-sound-database.

[2] SB. Yu, KW. Park, SH. Park., "Design and Implements of AI PCG Analyzer," *Conference on Institute of Embedded Engineering of Korea (IEMEK)*, 2019.

[3] S. Chatterjee, M. Rahman, E. Nemanti, J. Kuang, "WheezeD: Respiration Phase Based Wheeze Detection Using Acoustic Data from Pulmonary Patients Under Attack," *EAI International Conference on Pervasive Computing Technologies for Healthcare*, 2019.

KSII The 12th International Conference on Internet (ICONI) 2020.

Korean Fingerspelling Recognition from a Video via Attention Mechanism

Yoonyoung Jeong and Han-Mu Park*
Artificial Intelligence Research Center, Korea Electronics Technology Institute(KETI),
Seongnam, Republic of Korea
[e-mail: yoonyj@keti.re.kr, hanmu@keti.re.kr]
*Corresponding author: Han-Mu Park

Abstract

In this paper, we propose a Korean fingerspelling recognition framework that adopts attention mechanisms to analyze importance among video frames. The proposed framework provides a character-level recognition method that generates a sequence of Hangul from a Korean fingerspelling video. The method consists of a sampling step, which selects several frames from the video, a keypoint extraction step, which localizes arm and hand keypoints by using the OpenPose API, a recognition step, which utilizes positions of the keypoints as an input sequence and estimates each phoneme by using Seq2Seq and Transformer models. We have evaluated our framework on the Korean sign language dataset, and have shown meaningful results.

***Keywords*:** Fingerspelling, Deep Learning, Human pose estimation

1. Introduction

Fingerspelling is an expression system that represents each letter by using a series of hand gestures. Fingerspelling is very important for communication because sign language that is a mother tongue of hearing impaired people often have no appropriate sign words for representing novel proper nouns or foreign language. In the expression system, each letter is represented by using a series of hand gestures, and then each word is represented by combining the gestures.

Since fingerspelling has the character-level expression system, most previous researches considered the recognizing each letter from an image[1]. However, the recognition algorithm should consider image sequences because some letters in the Korean fingerspelling require hand movement. Another work[2] considered a video as an input, however it studied for English only. In this paper, we propose a framework that recognizes Korean fingerspelling from a video. The proposed framework can recognize each letter, Korean alphabet, from a sequence of images.

2. Korean fingerspelling recognition

The proposed framework consists of three steps: sampling step, keypoint extraction step, and recognition step.

2.1 Data preprocessing

To refine and augment a set of fingerspelling videos as the training/testing dataset, we propose two-step approach.
In the first step, sampling step, several appropriate frames are selected from an input video. Firstly, an interval is set, which indicates the motion for representing fingerspelling is started. Then, certain number of frames are randomly selected in the interval.

This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (2017-0-00255, Autonomous digital companion framework and application)

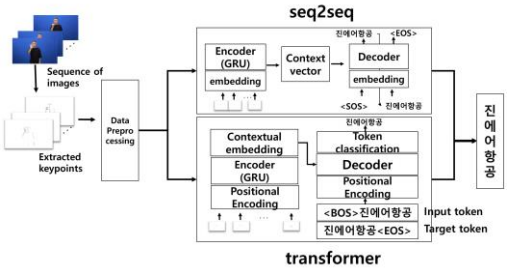


Fig. 1. Overall Architecture

In the next step, keypoint extraction step, the Openpose API[3] is applied to extract positions of hand joints. To overcome occlusion and blur effects, we also apply Kalman filter to track keypoints. Finally, the stabilized keypoints are collected to construct an input data for training and inference.

2.2 Recognition model

Fig. 1 represents overall recognition architecture of the proposed framework. To acquire meaningful information for recognizing fingerspelling, we only utilize right hand and arm(neck to wrist) keypoints. Then, the selected hand and arm keypoints are normalized by the length of middle finger bone and right shoulder bone respectively. Finally, the normalized positions of keypoints are used as an input vector of the deep neural networks. In this work, we adopt Seq2Seq[4] and transformer[5] models to analyze input vectors and generate Hangul. To apply Seq2Seq model, we utilized Gated Recurrent Units(GRUs) for each module of the encoder. To apply another case, transformer model, we utilized tokenizer, KoNLPy[6], then generated tokens are used as a series of input vectors and target sequences for decoder.

3. Experiments

3.1 Dataset

We have acquired videos of the 116 words that require for guiding the facilities in airports. We use KETI sign language fingerspelling video datasets[7] for training. There are 2,320 videos in total, consisting of 20 people at 60 FPS. Use 16 of them for training and the rest for testing.

3.2 Experimental results

Table 1 represents evaluation results according to the number of sampled frames per model. The seq2seq and transformer have shown the best performance at 30 and 40 frames respectively. The result means that the transformer model need more data than seq2seq for more accuracy.

Table 1. Result according to model and the number of sampled frames

Model	Sampled frames	BLUE-1
seq2seq	20	0.65
seq2seq	30	**0.66**
seq2seq	40	0.63
transformer	20	0.63
transformer	30	0.63
transformer	40	**0.65**

4. Conclusions

In this paper, we present a recognition method for Korean fingerspelling. The proposed framework has shown meaningful result for understanding Korean fingerspelling.

References

[1] S. Moon, Y. Kim, I. Suh, “Real-time Deep Convolutional Neural Network using Stereo Image for Finger language Recognition,” *The HCI Society of Korea*, 2017.

[2] B. Shi, A. M. Del Rio, J. Keane et al., “Fingerspelling recognition in the wild with iterative visual attention,” In *ICCV*, 2019.

[3] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh, “Openpose: Realtime multi-person 2d pose estimation using part affinity fields,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[4] I. Sutskever, O. Vinyals, Q. V.Le, "Sequence to Sequence Learning with Neural Networks," In *Advances in Neural Information Processing Systems*, 2014.

[5] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. "Attention is all you need," In *Neural Information Processing Systems (NIPS)*, 2017.

[6] E. L. Park and S. Cho, “Konlpy: Korean natural language processing in python,” In *Proceedings of the 26th Annual Conference on Human and Cognitive Language Technology*, Chuncheon, Korea, 2014.

[7] S.-K. Ko, C. J. Kim, H. Jung, and C. Cho, “Neural Sign Language Translation Based on Human Keypoint Estimation,” *Applied Sciences*, vol. 9, no. 13, pp. 2683, 2019.

KSII The 12th International Conference on Internet (ICONI) 2020.

Motion Transfer in Sign Language Using Unsupervised Learning

Boeun Kim, Yoonyoung Jeong, and Hyedong Jung
Artificial Intelligence Research Center, Korea Electronics Technology Institute, South Korea
[e-mail: kbe36, yoonyj, hudson@keti.re.kr]

Abstract

Motion transfer is a problem of generating a video in which the source object depicting the motion of the target video. In this paper, we propose a network that gets a single image of a person as input and transfers sign language motions of a target video that is captured in a standardized environment. The purpose is to create a model that generates a video with only a single image that is taken in general clothing and background so that the model can be used to actual applications. To tackle these issues, we leverage both standardized sign language videos and videos with arbitrary movements in various environments as training data. In addition, end-to-end learning using unlabeled data induces to extract key-points according to the motion characteristics of the training videos.

***Keywords*:** Motion transfer,, Image animation, Motion re-targeting, Unsupervised learning, Sign language

1. Introduction

In the communication system between the hearing impaired and the non-disabled, it is important for the hearing impaired to understand what the partner wants to say. Many hearing impaired do not have formal education to read texts, and those who communicate mainly in sign language (SL) rarely use texts. If they look at SL instead of texts, they much better understand the meaning. In this paper, the motion transfer method is applied in order to automatically generate videos with SL motions that non-disabled people wants to say.

Motion transfer, which is also called image animation or motion re-targeting, is a problem that makes the person of the source image imitate the motion of the target video. In this paper, by injecting the single source image of a speaker and a SL video of an expert, we create the video in which the speaker performs the same SL motion as the expert. We used Korean sign language datasets [1] that were recorded in a standardized clothing and environment.

2. Related Works

In recent studies, keypoint-based motion transfer methods show good performance. Chan et al. extracted the body keypoints of the source and target through pre-trained model and reconstructed the image using them [2]. Siarohin et al. introduced unsupervised learning to extract keypoint from arbitrary objects [3]. The method of Kim et al. [4] is related to video prediction, and they trained the keypoint detector and translator in end-to-end manner. The error was reduced by using the background masking and stat-of-the-art performance was shown.

The network of Kim et al. [4] was adopted to our motion translation task. In Kim's paper [4], the target video and the source image were extracted from the same data distribution, but in our task, since the standardized environment in

This research was supported IITP [2017-0-00255] and NIPA [S1601-20-1034] grant funded by the Korea government.

the SL video is different from the real world environment, we add newly captured training data.

3. Methods and Results

The model architecture is described in **Fig. 1 (a)**. Inspired by [4], we leverage U-Net as a keypoint detector and an CNN-based encoder-decoder structure as a translator. A part of the Korean sign language dataset [1] which is a video dataset of experts in black T-shirts in front of a blue screen is used for the target video. A lot of artifacts occur when a frame sequence is generated by inputting an image captured in a general environment into a network learned only with datasets recorded in a formal environment.

In this paper, newly captured datasets are additionally learned, and it is created by taking videos of a person moving her/his arm while standing. Since this model uses an unsupervised learning method, training data can be added simply without keypoint labeling. Total 305 SL video clips, 105 sentences recorded by three people, were used, and the length is at least 58 frames and up to 234 frames. The additional videos were recorded one person wearing three different costumes in two different places and divided into 200 clips of 100 frames long.

The results are illustrated in **Fig. 1. (b)-(e)**. It worked without artifacts in the demonstration in a general environment. Keypoints are automatically captured mainly in the arm parts which have a lot of movement in SL motions.

4. Conclusions and Discussions

It is verified that the motion transfer performs robustly in various environments by additionally learning the data captured in simple methods without labeling. During the test, it is possible to generate a video without an additional learning process for the images or videos of the source person.

When the posture of the person in the given source image is different from the posture of the person in the target video start frame, it works well without any problems. However, if their body proportions are different or they are not aligned, it causes many artifacts. To solve this problem, the normalization method should be

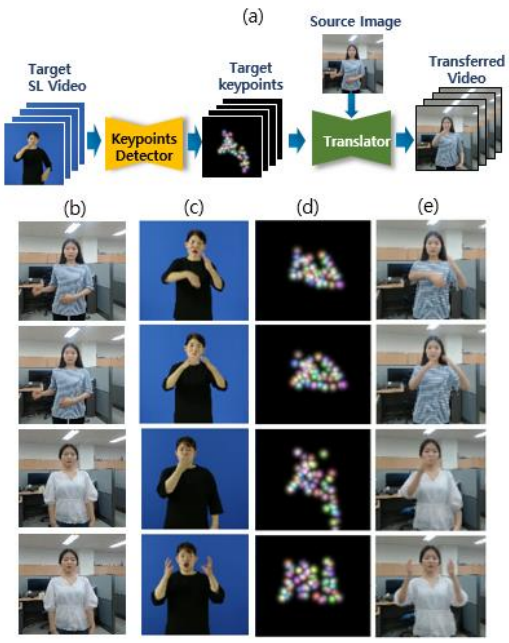


Fig. 1. (a) A model architecture inspired by [5]. (b) Source images, (c) Target video frames, (d) Extracted keypoints, (e) Transferred frames.

studied in future works. If the target object does not align with the source, there is a problem of restoring the occlusion area behind the person. This problem is expected to be alleviated when a large number of datasets are used for training.

References

[1] "Korean sign language dataset for vision basis AI development", *AI hub*, https://www.aihub.or.kr/aidata/7965

[2] Chan, Caroline, et al. "Everybody dance now." *Proceedings of the IEEE International Conference on Computer Vision*. 2019.

[3] Siarohin, Aliaksandr, et al. "Animating arbitrary objects via deep motion transfer." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019.

[4] Kim, Yunji, et al. "Unsupervised Keypoint Learning for Guiding Class-Conditional Video Prediction." *Advances in Neural Information Processing Systems*. 2019.

Dialogue Strategy Prediction Using Multi-modal Information for Empathic Dialogue Generation

Jin Yea Jang, Minyoung Jung, Hyedong Jung, and Saim Shin
Artificial Intelligence Research Center, KETI
Sungnam-si, South Korea
[{jinyea.jang, minyoung.jung, hudson, sishin}@keti.re.kr]
*Corresponding author : Saim Shin

Abstract

With the development of deep neural network technology and the high availability of big data, research on dialogue models that automatically generate dialogue sentences is being actively conducted. Research has also been paying great attention to creating more natural dialogue by reflecting various aspects of dialogue through the use of multi-modal information. This paper focuses on the prediction of dialogue strategies that are important for the generation of empathic response. We implemented a classifier that predicts the dialogue strategy for the next response based on the dialogue context as well as the information of emotions associated with multiple modalities (e.g. text or sound). The experimental results showed that our model helped predict dialogue strategies for the next response with n performance improvement (8%) than the baseline model. The results also showed that the information from text contributed more to predicting dialogue strategies for the next response than that from multimodal.

***Keywords*:** Artificial intelligence, Empathic response generation, Multimodality

1. Introduction

Recently, much interest and attention has been paid to the development of AI-based empathic response systems that can help people with symptoms of depression [1]. In order to develop a dialogue model for empathic responses, it is important to utilize emotion information from various modalities, such as text, sound, because these types of sensory information are associated with emotions [2]. In this paper, we present a model that predicts dialogue strategies by using user utterance and emotion information from text and multimodal to generate empathic responses. The paper proceeds in the order of data construction, introduction of the prediction model, and experiment results.

2. Experiments

2.1 Empathic Dialogue Data

Table 1. Examples of the system's responsive strategies (Italic: translated into English from Korean)

Strategy	Example
Clarification	*Are you in a sad mood now?*
Empathic response	*It must be difficult.*
Suggestion	*Why don't you go for a walk?*

We design the dataset based on three major empathic dialogue strategies by utilizing the actual psychological counseling record data and

This study was supported by Institute for Information & Communications Technology Promotion (IITP) and National IT Industry Promotion Agency (NIPA) grant funded by the Korea government (2017-0-00255 and S1601-20-1034).

the advice of two psychiatrists. **Table 1** shows examples of system responses based on different strategies.

Besides of the strategies, we also tagged seven emotional labels (Ekman's six basic emotions [3]: anger, sadness, fear/anxiety, surprise, happiness, disgust; plus neutral) for the text and multimodal for each utterance of the dialogue. The emotion labels for multimodal here is a comprehensive emotion that can be inferred from not only the text of the utterance, but also the sound or facial expression. A total of 4,915 dialogue pairs (between a user and an empathic dialogue system) were constructed, and the distribution of emotion information for text and multimodal is shown in **Fig. 1**. The emotional label according to the modality may be different for the same dialogue, which means that there is a possibility to influence the dialogue strategy prediction depending on how we use the information from different modalities.

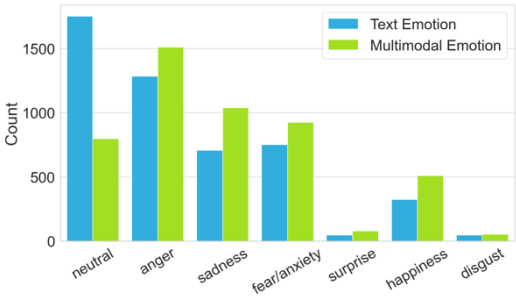


Fig. 1. Distribution of emotion labels from the text and multimodal information

2.2 Prediction Model Development

We used KoBERT [4], a pre-trained language model learned with Korean data. The overall model architecture is shown in **Fig. 2**. The input of the prediction model is composed of emotion and utterance information. We also conducted an experiment on the influence of the use of the history information in the previous dialogue on strategy prediction. The history information was additionally added to the utterance information.

The prediction model computes representation vectors, and the first representation vector passes through the linear and the softmax layer and outputs a probability value for each strategy.

The fine-tuning of the model was done using the data we built. The data were divided into train, validation, and test sets at a ratio of 3:1:1. The optimizer used for training was AdamW [5], the epoch was set to 10, and the learning rate was 0.00002.

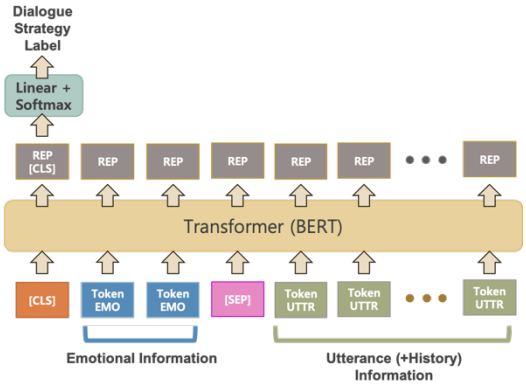


Fig. 2. Architecture of the dialogue strategy prediction model

2.3 Results

We trained different models based on different modalities and the use of history information as follows.

- Uttr only is a baseline without emotions.
- Uttr+T.Emo is a model with text emotions.
- Uttr+M.Emo is a model with multimodal emotions.
- Uttr+H is a model with dialogue history.
- Uttr+H+T.Emo is a model with dialogue history and text emotions.
- Uttr+H+M.Emo is a model with dialogue history and multimodal emotions.

We measured the F1 scores through 5-fold cross-validation. **Table 2** shows the results.

Table. 2. Average F1 score for each strategy (CL: Clarification, EM: Empathic response, SG: Suggestion) and total depending on the combination of input information

Model	CL	EM	SG	Total
Uttr only (Baseline)	0.53	0.57	0.76	0.63
Uttr+T.Emo	0.61	0.63	0.75	0.67
Uttr+M.Emo	0.59	0.59	0.74	0.64
Uttr+H	0.62	0.54	0.83	0.67
Uttr+H+T.Emo	**0.67**	0.62	**0.84**	**0.71**
Uttr+H+M.Emo	0.66	**0.65**	**0.84**	**0.71**

The results show that the models with emotion information have the better predictive performance than the models with only utterance information. In addition, the uses of dialogue history help improve model performance. In the

comparison of the modality of emotion information, there was a difference in performance. The use of text emotion information showed slightly better performance than the use of multimodal information.

3. Conclusions

We presented that using emotion information in dialogue can help predict the dialogue strategy. Our future work is to extend our model with more responsive strategies as well as develop a model that generates an empathic response.

References

[1] P. Kataria, K. Road, A, Jain, P. Dwivedi, and S. Bhingakar, "User adaptive chatbot for mitigating depression," *International Journal of Pure and Applied Mathematics*, Vol 118, No.16, pp.349-361.

[2] Z. Zhang, L. Liao, M. Huang, X. Zhu, and T. Chua, "Neural multimodal belief tracker with adaptive attention for dialogue systems." In The World Wide Web Conference, pp. 2401-2412. 2019.

[3] P. Ekman, "An argument for basic emotions," *Cognition & Emotion*, Vol 6, No. 3, pp. 169-200, 1992.

[4] https://github.com/SKTBrain/KoBERT

[5] D. P. Kingma, and J Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 2014.

Allergy Alarm System based on Food Content using Object Recognition

[1]Kwang Ok Lee, [2]Seo Won Chang, [2]Seo In Lim, [3]Keun Chang Kwak, and [1*]In Seop Na
[1]National Program of Excellence in Software Center, Chosun University, Gwangju, South Korea
[2]Department of Computer Science, Chosun University, Gwangju, South Korea
[3]Department of Electronics Engineering, Chosun University, Gwangju, South Korea
[e-mail: csskwang@naver.com, cseowon@gmail.com, dlatjdls0112@naver.com, kwak@chosun.ac.kr, ypencil@hanmail.net]
*Corresponding author: In Seop Na

Abstract

There is lack information about food ingredients or food allergies that caused allergic shock in infants and toddlers. In Europe and the United States, it is compulsory for food allergy and ingredient information to be notified on the product packaging, allowing users to easily verify the information. In this paper, we proposed that the image data-based video classification using tensorflow which is implemented by putting the photos of foods in circulation on the market. We want to take a picture of a food product with a camera and immediately collect information to quickly identify whether you are allergic to the product and the type of food additive.

***Keywords*:** Allergy, Tensorflow, CNN, Convolution, Food, Object Recognition

1. Introduction

Korea lacks a lot of allergies, food information and knowledge compared to developed countries, and in order to check the information you want, you have to check the ingredients on the back side every time you buy food[2]. If you take just one photo of the additive and allergy-causing element of processed food, you can easily check the information of the product you want to check in a short time[3].

This can lower the incidence of shock patients from allergic reactions and raise social awareness and interest in allergies. It also supports languages from various countries to help people around the world identify nutritional information and allergy types of Korean food to prevent risks and to help people on certain diets such as vegetarianism, halal and kosher.

Therefore, we try to provide information on processed foods to reconsider the social statement of allergies and indifference. It implements data-based video classification CNN using tensorflow by putting photos and information of processed foods on the market into data sets. Currently, similar products are 'Alergy Tell Me' and 'Homseon' in Korea[4,5].

In order to collect information such as photos and voice recognition, we will collect information quickly through product photography, product recognition, food allergy status, and food additives so that people with diseases and allergies can conveniently search for information.

2. Proposed Method

2.1 Data set

The data set first selects 20 processed foods(chickensoup, breast, duck, jaban,peachcan, tofusushi, ramen, tofu, soup, curry, dumpling, fishcake, hotdog, laver, pancake, pasta, rice, spam, tuna, udong) and collects a total of more

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (NRF-2018R1A1A1A05022526) and (NRF-2019K2A9A1A06100184)

than 10,000 data. Second, it collects images of processed food through 20 kinds of googling. There is a limit to collecting public data, so you buy products to collect your own datasets. We collect pictures from various angles (environmental) because they need to be recognized when taking pictures with mobile phones.

Fig. 1. Example of Udong Data set

The nutrients of the selected 20 processed foods are obtained from the Food and Drug Safety Administration's Food and Nutrition Database and Food Safety Country, and allergy and cross-contamination information can be obtained from the back of the food.

2.2 Convolution Layer

The Convolution Layer consists of a composite product, channel, filter, padding, etc. The channel usually consists of three RGB channels (black and white images consist of two-dimensional data consisting of one channel) For our model, the coli shape is 28 pixels high and 28 pixels wide, and the black and white image is (28, 28, 1).

As shown in **Fig. 2**, a filter is a matrix for extracting the characteristics of an image. Usually defined as a 3*3 or 4*4 square matrix, and if the data are characteristic, it is a large value, if not close to zero. The process of figuring out what the value matrix (filter) is for finding features is a learning process for CNN. The stride is the interval at which the filter moves, and if stride=1, it moves one space each and performs a composite product of input data and filters.

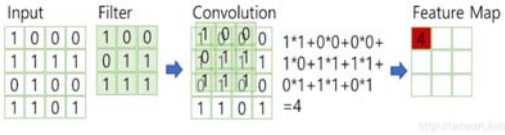


Fig. 2. Convolution Filter

Doing so will result in the final output of the transformation layer being smaller than the initial input, and eventually results will be lost before the neural network is applied and even before extracting a large amount of data and sufficient characteristic values. Therefore, to make the output size of the conversion layer equal to the input data size, the data is populated with specific values (mainly 0) outside the data, which is called padding.

It is consists of a total of three layers and applies the relu and SoftMax functions as the activation functions. It also applies dropout to prevent over-conformity. Reduce data space using the pooling layer. In the composite layer, the output data is sized to the input data size and resizes only in the pooling layer.

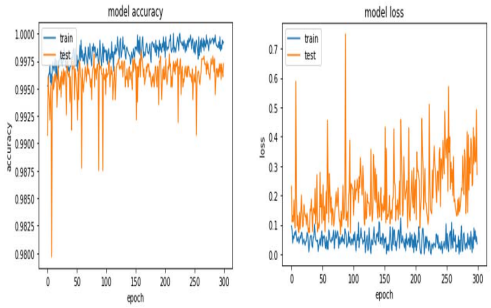


Fig. 3. Accuracy and Loss

2.3 Data Augmenting

Before learning using a generator, check the modified image for abnormalities, move the rotation range of the image parallel to 40 degrees, 0.2%, and then expand it to 1/255 to switch to the 0-1 range. In addition, 600 is designated for each model because balanced learning is difficult if the amount of learning data is not constant for each category. To learn about image files, save the image files in a numerical array and then resize the image to 28*28 to reduce the data capacity. Then, divide the learning by 256 because it requires less than a decimal point of 0 and 1. -> X.append(img/256)

2.4 Learning model

CNN(Convolution Neural Network) is an artificial neural network algorithm called synthetic-square neural network. This algorithm was first introduced in 1998, but has been used in so many areas in the field of image recognition.

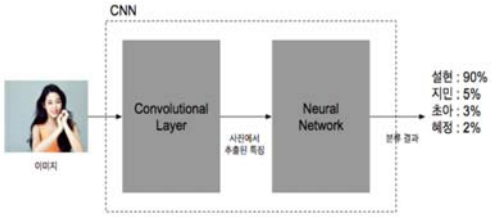


Fig. 4. Validating the CNN Model

CNN is like having multiple layers of Convolution Layers in front of a basic neural network.

The above transition layer part becomes the characteristic extension part of the picture below, and the neural network part becomes the classification part of the picture below. After the combination layer and pooling layer have been stacked together to fully extract the properties, the flat layer is used to form the array and finally create an array-type full-connect layer and apply it to the classification model. The activation map with activation function is the final output of the conversion layer generated by the following procedure.

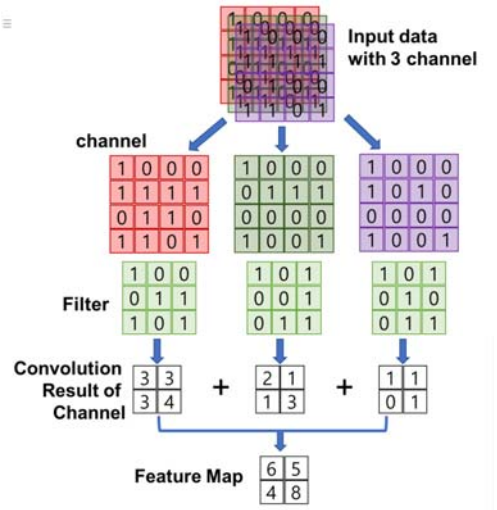


Fig. 5. Output Results for Generated Conversion Layers

4. Conclusions

Since Korea lacks a lot of allergies and knowledge of food information, it is necessary to check the ingredients in the back whenever it buys food. However, it is easy to check the information of the products you want to check in a short time by taking only one photo of the additives and allergy-causing elements of processed foods.

In this paper, data-based video classification using tensorflow which was implemented by putting photos and information of processed foods on the market in a data set, allowing people with diseases and allergies to quickly gather information by taking product photos and identifying whether they are allergic to food and types of food additives. This can lower the incidence of shock patients from allergic reactions and raise social awareness and interest in allergies. It can also support languages from various countries to help people around the world identify nutritional information and allergy types of Korean food to prevent risks and to help people with certain diets, such as vegetarian, halal and kosher.

References

[1] Mi Jung Kim, Ki Hyeon Sim, “Study on Incidence and Improvement of Food Allergies for Prevention of Damage in Adolescents and Adults” *The East Asian Society of Dietary Life*, Vol 24, No. 4, pp. 514-528, 2014.

[2] RS. Gupta, EE. Springston, B. Smith, Js. Kim, JA. Pongracic, X. Wang, J. Holl, “Food allergy knowledge, attitudes, and beliefs of parents with food-allergic children in the United States”, *Pediatr Allergy Immunol 21*, pp. 927-934, 2010.

[3] M. Groetch, H. Sampson, “Management of food allergy”, *Principles and Practice 2nd ed Philadelphia : Saunders*. pp. 541-551, 2010.

[4] Change the standard of safe and enjoyable food experience. https://www.facebook.com/allergyi

[5] Where consumers gather to wait for new food every day. http://www.umsun.co.kr/

Lip Shape based Deep Learning for Pronunciation Recognition

[1]Sang Ok Youn, [2]Hee Won Kim, [2]Tae Hee Kim and [1*]In Seop Na
[1]National Program of Excellence in Software Center, Chosun University, Gwangju, South Korea
[2]Department of Computer Science, Chosun University, Gwangju, South Korea
[e-mail: ysangok@chosun.ac.kr, h-won@naver.com,grandejoie15@naver.com, ypencil@hanmail.net]
*Corresponding author: In Seop Na

Abstract

In general, it is very difficult for people to determine whether they are good pronunciation when they see a letter and pronounce it. There is a limitation to studying English pronunciation by not navtive speaker like a Korean English teacher in Korea. According to the principle of pronunciation, it is very important factors which is the shape of the lips and position of the tongue. In this paper, we present the architecture of recognizing the shape of the lips and the exacting the position of the tongue to know the correct pronounce or not. At first, we use the YOLO to detect the shape of lips. Secondly, we extract the lips based on facial landmark. To train the shape of lips, as possible as we gather the shape of lips for the same pronounce in English news video and English lecture video by googling. When make the dataset, only we are forcusing the less number of the 1 of the loss value. After detect the lips position, we visualize the ploygon on the image. Accoding to the national standard landmark, we define the position of lips for the lips' tail, lip's upper center, lip's lower center and so on. And also we categorize the lips by gender and ages.

Keywords*:* Pronunciation, Lip Shape, Deep Learning, YOLO, English

1. Introduction

Most of the current functions of the pronounciation related applications do not provide the pronunciation position for better understanding the pronunciation or to mimic the pronunciation. In this study, we proposes a new

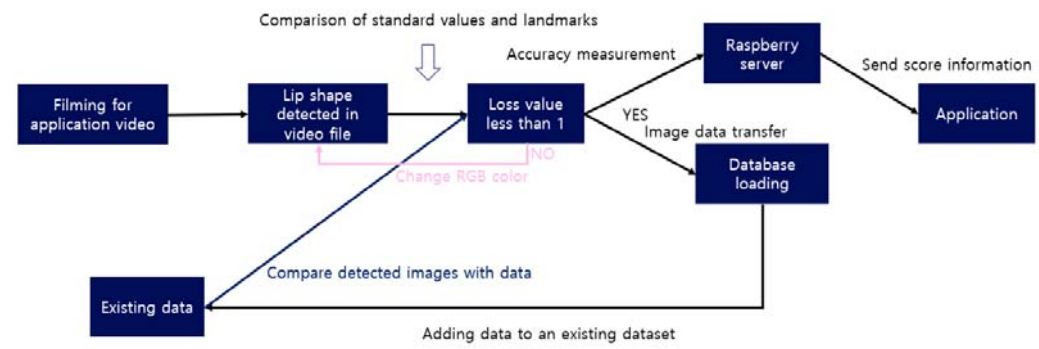


Fig. 1. The Flowchart of Proposed Method

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (NRF-2018R1A1A1A05022526) and (NRF-2019K2A9A1A06100184)
.

method for correcting pronunciation through image recognition in order to train the foreign language especially English speaking.
For the traing the shape of lips, we implement the gathering application for the shape of lips video. It detect and segment the shape of lips in real time by uploading to the server. One side we gather the dataset and another side we compare the pronunciation based on pre-trained dataset. After the comparision, we correct the shape of lips.

2. Proposed Method

As shown in **Fig. 1,** first step is capturing the video clip from camcoder or smart phone's camera, what we called the Filming Step, second step are detecting the lip shape and recognizing the lip points from facial landmark, third step is estimating the loss value as accuracy measurement. Fourth step is sending the accuracy measurement to Rasberry server and sending the image datato the database at the same time. In final step, Rasberry server make a decision.
In **Fig.2,** we extract the facial landmark to detect the lip shape. After that we measure the loss value of lip shape to recognize the pronumciation related lip shape.
We use the LSTM for the learner's vocalization and classified by phoneme through softmax like **Fig3.** The YOLO algorithm divides the original image into equally sized grids, estimates the number of bounding boxes specified with a predefined shape around the center of the grid for each grid, and calculates the confidence based on this.

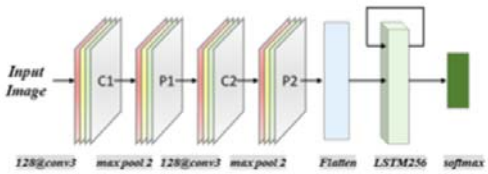


Fig. 3. Pipeline of vocalization

To detect the specified location and size of object, strategically it is chosen by scenarios to divide the original image into fixed-sized grid areas.
The algorithm predicts a fixed number of objects of a predetermined shape and size for each area by real-time detection.

3. Experiment and Results

We gather the voice and various lips shape dataset by our selves as shown in **Fig. 4.**

Fig. 4. Example of the lip shape data

In **Fig. 5.** shows that examples of the lips shape detection from our own camera and drama video clips. We marked the lip area based on national standard landmark position information.

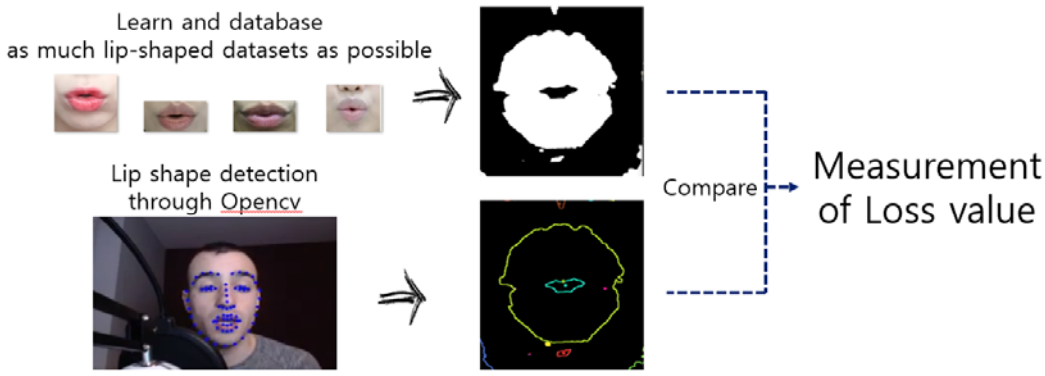


Fig. 2. Pipe line of recognition of lip shape

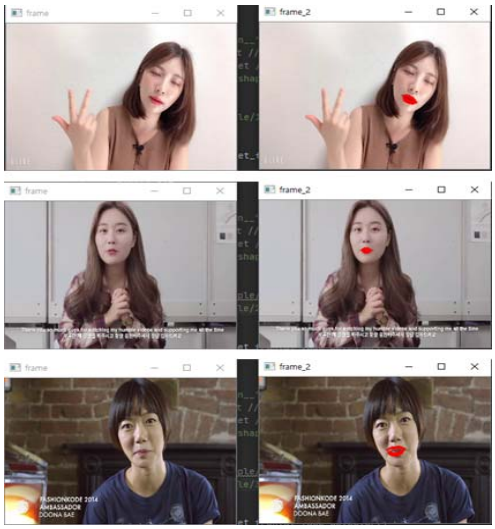


Fig. 5. Example of the detected lip shape area

In **Fig. 6** shows that loss information related with iteration of training.

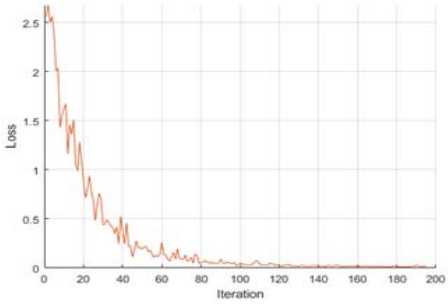


Fig . 6. Example of the lip shape data

4. Conclusions

In order to realize accurate pronunciation by recognizing the exact position of the lips and tongue without focusing on simple vocalization and speaking training, video recognition can be implemented to recognize the shape of the lips, uploaded to the server, and compared with the learned contents to correct the pronunciation. I studied about the APP that is there.

As a result, accurate pronunciation was possible, and it is thought that it can be applied to foreign languages other than English through the expansion of functions.

References

[1] S. A. Seshia, "Compositional verification without compositional specification for learning-based systems," University of California at Berkeley, pp. 1-8, Nov, 2017.

[2] A. Dimou, P. Medentzidou, F. Alvarez Garc´ıa, and P. Daras, "Multi-target detection in cctv footage for tracking applications using deep learning techniques," IEEE, ICIP 2016, pp. 928-932,

[3] X. Yin, W. Shen, J. Sarnarabandu, and X. Wang, "Human activity detection based on multiple smart phone sensors and machine learning algorithms," IEEE, CSCWD, pp. 582- 587, May 2015.

[4] X. Zhu and D. Ramanan, "Face detection, pose estimation, and landmark localization in the wild," IEEE, pp. 2879- 2886, 2012.

[5] X. Yu, J. Huang, S. Zhang, W. Yan, and D. N. Metaxas, "Pose-free facial landmark fitting via optimized part mixtures and cascaded deformable shape model," ICCV 2013, pp. 1944-1951, Dec. 2013.

[6] H. Boukamcha, M. Hallek, F. Smach, and M. Atri, "Automatic landmark detection and 3D Face data extraction," ELSEVIER, Journal of Computational Science 21, pp. 340- 348, Nov. 2016.

[7] X. P. Burgos-Artizzu, P. Perona, and P. Doll´ar, "Robust face landmark estimation under occlusion," ICCV 2016, pp. 1513-1520, Jan. 2016.

[8] S. Ren, X. Cao, Y. Wei, and J. Sun, "Face alignment at 3000 FPS via regressing local binary features," CVPR 2014, pp. 1-8, Jun. 2014.

[9] K. Zhang, Z. Zhang, and Z. Li, "Joint face detection and alignment using multitask cascaded convolutional networks," IEEE, Signal Processing Letters, vol. 23, no. 10, pp. 1499-1503, Oct. 2016.

[10] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "300 Faces In-The-Wild Challenge: database and results," ELSEVIER, Image and Vision Computing 47, pp. 3-18, Jan. 2016.

[11] V. Kazemi and J. Sullivan, "One Millisecond Face Alignment with an Ensemble of Regression Trees," CVPR 2014, pp. 1-8, Jun. 2014.

The Shortest Path Finding using Reinforcement Learning under Labyrinth

[1]Sung Ki Yang, [2]Mok Gyun Jo, [2]Hwi Hwan Jung, [1*]In Seop Na
[1]National Program of Excellence in Software Center, Chosun University, Gwangju, South Korea
[2]Department of Computer Science, Chosun University, Gwangju, South Korea
[e-mail: yskstudy@chosun.ac.kr, ahrrbs1@naver.com, gnlghks1@naver.com, ypencil@hanmail.net]
*Corresponding author: In Seop Na

Abstract

In this paper, we study the shortest paths finding problem such on the paths, obstacles, and walls using artificial intelligence techniques in a given fixed maze. The learning method to find the shortest path was composed of Python's Tkinter, and the shortest path was learned through reinforced learning using Q-learning. There is no set time to identify the maze, and the average number of learning before learning is about 20, and when learning is completed, the learning time to the destination may vary depending on Lr.

Keywords*:* Artificial intelligence, Reinforcement learning, Q-Learning, Path finding, Maze

1. Introduction

With the 4th Industrial Revolution, the AI market is growing rapidly as companies invest more in AI technology every year. In field of AI, optimization problem is traditional regacy work but it is core algorithm to solve the AI problem. That is why, many researchers are trying to solve the shortest path under the maze.
Labyrinth[1,2], as shown in Fig. 1. , is being studied in finding the shortest route through self-learning, and research on robot vacuum cleaners, autonomous vehicles, and drones is also rapidly becoming established.

Fig. 1. A Example of Labyrinth, source: wikipedia

The proposed method is the shortest path finding learning program using artificial intelligence. Instead of trying to escape the maze with a keyboard, it identifies the maze structure through Q-function based reinforcement learning and displays the number of travels and arrival time after moving to the destination by itself. .

2. Shortest Path Fining Model

The proposed pipeline of implementation is shown in **Fig. 2**. Reinforcement learning [3] can learn optimized path search without information of specific rules for a given target. It is a learning program to find a route that saves and searches the location of a random maze, unexpected unexpected situations in the middle, and situations with multiple destinations in the process of finding a route and going to a destination. The learning procedure is as follows.

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (NRF-2018R1A1A1A05022526) and (NRF-2019K2A9A1A06100184)

Fig. 2. Proposed method

1 *Generate the Maze: Labyrinth Algorithm Implementation by python which is a widely used language for machine learning*
2. *Draw the Maze Map: Create maze map by TKinter which is the module for GUI in Python*
3. *Finding a shortest path through Reinforcement learning which is a treadmill that can escape the maze by itself through Q-Learning learning.*

Generated Maze System Diagram is shown in **Fig. 3.** In this maze, it grasps the maze structure through Q-learning based reinforcement learning rather than attempting to escape the maze with a keyboard and displays the number of movements and the time to reach after moving to the destination by itself.

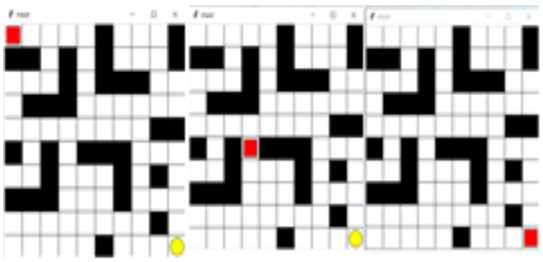
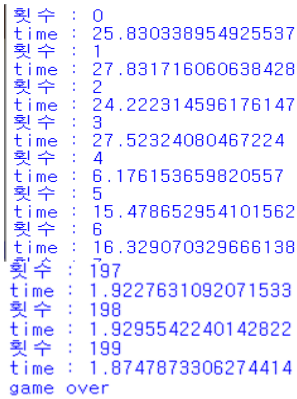


Fig. 3. A Example of Generated Maze System Diagram

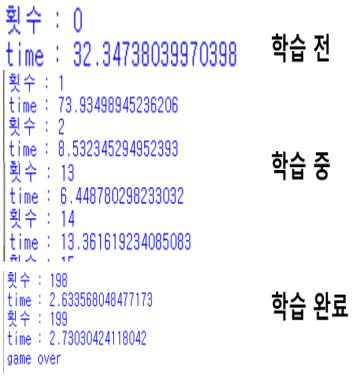
3. Experiment Result

The time or number of times until the first learning is completed is similar, but the time it takes to reach the destination decreases if the learning rate value is larger as the learning is repeated. When the Lr(Learning rate) value is large, the number of times to learn to the destination decreases, when the learning is completed, the time to the destination decreases, and when the Lr value is small, the number of times to learn to the destination increases, and when the learning is completed, the time to the destination increases.

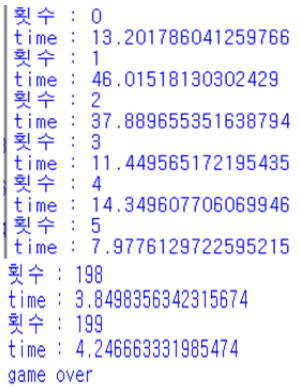
In **Fig. 4a, Fig. 4b, Fig. 4c** are show that the proposed experimental results (before learning vs. in learning vs. learning completed)



(a) Lr : 1.0



(b) Lr : 0.8



(c) Lr : 0.6

Fig. 4. Comparision of Learning rate

4. Conclusions

The shortest path finding learning program using artificial intelligence allows you to set up and explore virtual terrain in advance when exploring caves or remote areas, and paramedics can quickly perform lifesaving in industrial and emergency situations. In everyday life, the structure of the inside and outside of the building was identified so that the work could be performed quickly.

References

[1] M.Vela-Pérez, M.A.Fontelos, J.J.L. Velázquez, “Ant foraging and geodesic paths in labyrinths: Analytical and computational results,” *Journal of Theoretical Biology*, Vol. 320, No. 7, pp. 100-112, 2013.

[2] L.L. Cristea, B. Steinsky, “Curves of infinite length in 4×4-labyrinth fractals,” *Geom. Dedicata*, Vol. 141, No. 1, pp. 1-17, 2009.

[3] Mnih, Volodymyr; et al., "Human-level control through deep reinforcement learning". *Nature*, Vol. 518, No. 7540, pp. 529–533, 2015.

[4] Won Ouk Kim, Dae Hee Kim, Dae Gwun Youn, “Optimum Evacuation Route Calculation Using AI Q-Learning, ” *in The Korean Society Of Marine Environment & Safety*, 2018.

[5] Tae kyung Lee, Jun Ri Jeon, “Destination Path Search using Reinforcement Learning Technique,“ Proceedings of the Korea Information Processing Society Conference, 2013.

[6] Se Il Lee, “Dynamic Programming Algori thm Path-finding for Applying Game,” *Journal of the Korea Society of Computer and Information*, Vol.10, No. 4, pp. 213-219, 2005.

[7] Mee Kyung Nam, “A Study on Present Condition of Development and Market of Artificial Intelligence Robot In and Out of Country,” *The Korean Society Of Design Culture* , Vol. 16, No. 2, pp.198-207, 2010.

[8] Hyang Joon Yoo, Seung Su Mok, Jun Seo Kim et al, “Control Technology Based on the Finger Recognition of Robot Cleaners,” *The Journal of the Korea Institute of Electronic Communication Sciences*, Vol.15 No.1, 2020.

[9] Seung Ju Jang, “Autonomous vehicle related SW technology trend,” *Information & communications magazine*, Vol.33, pp.27-33, 2016.

Security and Privacy in Big Data: A review of current issues and preserving techniques

Van Y Pham
Discipline of Information & Communication Technology, University of Tasmania, Australia
[email: vypham@utas.edu.au]

Soonja Yeom*
Discipline of Information & Communication Technology, University of Tasmania, Australia
[email: soonja.yeom@utas.edu.au]

In Seop Na
National Program of Excellence in Software Centre, Chosun University, South Korea
[email: ypencil@hanmail.net]

*Corresponding Author: Soonja Yeom

Abstract

Big Data, the term refers to an enormous amount of a variety of data, has brought many benefits to organizations and governments with the results found from data processing. However, there remains a significant risk for many users to be personal exposure data which arise from the process of storing, managing and analyzing Big Data. Moreover, the challenges arise from privacy concerns of using data as well as data collection processes. In this paper, we argue approaches of Big Data security and privacy, then review current issues and preserving techniques in Big Data context. Finally, we address challenges to present some comprehensive protection solutions that enhance data security and privacy in a Big Data scenario.

Keywords: Big Data, big data security, big data privacy, big data issues, big data preserving, security, privacy.

1. Introduction

The development of cutting-edge technologies these days, which is significantly increasing in many aspects such as social media, economics, environment, education and research, has resulted in the collection of data. These kinds of data are commonly known as "Big Data". Big Data (BD) is no longer a new research topic because it has been studied for couples of decades, for example, in social and computing sciences [1, 2]. Analyzing BD to expose useful patterns can enhance the process of decision-making not only in business but also in governance, research and academic. There are five beneficial ways that BD can bring to organizations: enhancing cybersecurity, increasing customer satisfaction, understanding user preferences through social media, detecting fraud transactions in internet banking, and identifying risks in the financial area [3]. However, BD also leads to many difficulties in collecting, analyzing, storing, managing and distributing enormous and complex sets of data [4]. The concern also encompasses security and privacy issues in BD.

This study is to identify and synthesize the current issues as well as preserving techniques in the context of BD security and privacy. Whereby, the literature review section is initially obtained, followed by the methodology where the main method of conducting the research is adopted. Then, the authors discuss the results of reviewing the extant literature and suggest reasonable solutions to protect data in BD environment before drawing the research conclusions.

2. Literature review

BD is defined in its structure and functionalities. Researchers exposed that BD is bound by four V's of data: volume, variety, velocity and veracity [5]. BD is also discussed as the term that "describes a phenomenon involving complex and dynamic growth in data" [7]. It refers to the new way of applying digital technologies to extract a huge amount of various kinds of data in a high-speed process to discover, analyse and evaluate data [8]. In the functional dimension, BD represents the use of digital technologies to capture, store, manage, analyze and distribute enormous sets of data [4]. These functionalities of big data have caused too many concerns and issues in terms of data security and privacy nowadays. Thus, BD pertains to the word that implies the complication of data about its quantity and quality.

Security is defined as methods, processes and technological evaluations that used to protect data from unauthorized access, theft, replacement or damages of systems and devices within four dimensions, including framework security, data protection, data administration, and veracity and reflection security [10, 11]. BD not only causes many security and privacy challenges, but it also creates new issues which need to be handled by new approaches [9]. Whilst privacy is defined as the "claim of individuals to be left alone, free from surveillance or interference from other individuals, systems or organizations" [10], researchers also revealed that privacy describes the management of aspects: physical, social, informational, psychological privacy, decision and disposition [12, 15]. In the context of BD, "privacy is sensitive information that the data owners do not want to be disclosed, such as the representation characteristics of sensitive data and data itself" [14]. Also, privacy contains two vital keys that are confidentiality and fair use [13]. For example, organizations nowadays encrypt users' emails and hide IP address from the tracking of web servers, or hide geographic location when customers use smartphones.

BD security and privacy issues have become the most controversial topic not only on social media such as Facebook, Twitter, but also in academic research [17-20]. BD issues refer to situations that all queries of encrypted BD are secured in cloud services [9], or the failure of protecting any data from harmful attacks, lost, stolen for profit [16]. The security and privacy issues of BD are also discussed under external and internal risk perspectives. An instance of the external risk aspect, hackers attack the cybersecurity system and penetrate a business's database to attain private information or confidential data [20-24]. An internal risk aspect of BD security and privacy issues can be viewed by the misadventures of Orbitz and Netflix when these companies applied the price discrimination on groups of customers that could be outrageous after the enterprises had dealt with users' data [25].

3. Methodology

The research aims to review existing problems and current techniques to protect data in the scenario of BD security and privacy. Hence, reviewing the extant studies related to the topic is the meaningful and vital way that not only contributes to the achievement of the study, but it also provides possible suggestions to the existing literature [26]. To do that, we adopted the *scoping review* method to indicate the nature of the current literature on the topic [27-29], review research activities in Big Data security and privacy, identify research gaps of the extant studies [30], and achieve the comprehensive point of this review paper [27]. At the same time, we applied the method of Templier and Paré [31], which consists of six following steps to review the existing works: 1 - establishing the research question(s) and objective(s); 2 - searching the current works; 3 - selecting studies by using inclusion criteria; 4 - evaluating the quality of studies; 5 - extracting and gathering data to capture what is mostly involved to the topic [32]; 6 - analyzing and synthesizing data. Moreover, the result of reviewing literature provides perspectives which make sense of existing knowledge [33] that was presented in conceptual papers, empirical studies, journal articles that investigated problems and introduced important suggestions [34]. Besides, the researchers used the inclusive ethnography approach to extract all relevant evidence from the extant studies [35, 36] combined with using the thematic technique to identify, categorize and create right patterns from the data gathered [37].

4. Literature review results

4.1 Current issues

The finding of issues is synthesized in **Table 1**, where the "Papers" column indicates the papers that reviewed or discussed particular problems and issues grouped based on the similar technical problems or similar entities. The set of papers noted in **Table 1** below is the results of reviewing the extant literature.

Table 1. Current issues in BD security and privacy

Issues	Papers
Governmental spy; unauthorized data collection of organizations; third parties (hackers)	[6]
Human intervention (data processing, surveillance, destructive)	[38], [43], [42]
Traditional penetrations (Trojans, malwares, spywares, denial of services)	[6]
Cloud services	[6], [40]
Authorization and accessibility	[6], [40]
Data theft	[6]
Data transformation between large-scale network	[39], [6]
Accountability and auditing	[40]
Confidentiality and anonymity	[38], [41], [42]
Technological framework (Hadoop, HDFS, MapReduce)	[40], [41]
Encryption	[40], [42]

Infrastructure security (availability, authentication, architecture)	[41]
Integrity	[41]
Legislation	[41]

4.2 Preserving techniques

All current preserving techniques for BD security and privacy are visualized in **Table 2**, which the set of papers noted is the results of reviewing the extant literature.

Table 2. Current preserving techniques in BD security and privacy

Techniques	Papers
Anonymization (identity based, k-based)	[14], [21], [41], [44], [47]
Fast anonymization of BD streams	[21]
Access control	[14], [41]
Encryption (homomorphic, attribute-based, normal)	[14], [40], [41], [47]
Digital watermarking	[14]
Data provenance	[14], [40]
Differential privacy-preserving	[14], [21], [41]
Secure multi-party computation	[14], [47]
MapReduce framework	[21], [41], [44]
Privacy-preserving BD publishing	[21]
Fraud/attack detection	[40], [41]
Auditing	[40], [41]
Confidentiality preserving	[40], [41]
Integrity preserving	[40], [41], [44]
Cryptography	[41], [44], [45]
Cosine similarity computing protocol	[44]
Secure storage & sharing cloud tenants	[44]
C4.5 algorithm	[44]
Perturbation	[44]
Generating sufficiency-based non-synthetic perturbed data	[44]
Hardware security	[45]

4.3 Challenges

The result of addressing current challenges is shown in **Table 3** (The set of papers noted is the results of reviewing the extant literature).

Table 3. Challenges in BD context

Challenges	Papers
Secure computing in distributed programming frameworks	[11], [14], [46], [40], [42], [47], [21]
Security optimal scheme for non-relation data storage	[11], [14], [46], [21]
Scalable and composable privacy-preserving data mining and analysis	[11], [14], [46], [40]
Mandatory encryption for data-centric security	[11], [14], [46], [42]
Granular access control	[11], [14], [46], [40]
Secure data storage and transaction logs	[11], [14], [46]
Granular audits / Security audits	[11], [14], [46], [40]
Data provenance	[11], [14], [46], [40]
End-point input validation/filtering	[11], [14], [46]
Real-time security monitoring	[11], [14], [46]
Fake data generation / Generation	[40], [21]
Cryptographic protection	[40]
NoSQL databases	[40]
Big volume-driven	[44], [45], [42]
Big velocity-driven	[44], [45]
Big variety-driven	[44], [45]
Big veracity-driven	[44], [45]
Data integrity	[45], [21]
Justification of Big Data economics	[42]
Changes of culture	[42]
Lack of talent and infrastructure technology limitation	[42]
Storage and management	[47], [21]

5. Discussion

Firstly, the current issues of security and privacy in the BD context are diversity and variety. **Table 1** above shows that most issues come from human intervention, which relates to data processing and analyzing, and also originate from confidentiality and anonymity problems. This means people who interact directly or indirectly with data, have made many mistakes in the security and privacy of data. The second top issues recorded are cloud services, authorization and accessibility, technological framework, encryption, and data transformation. The issues make sense of the nature of big data. The amount of data increase will require huge storages and consistency. Also, the higher volume of various data the machines process, the greater speed and framework they need. In contrast, the impacts of traditional penetrations, data theft, infrastructure, integrity and regulation system in the context of BD are not attractive for researchers to review or study. At the same time, the probability of the issues, which come from business and environmental factors, is likely to occur, but the extant works have not examined as well as reviewed these issues.

Secondly, there are many existing techniques for protecting BD, especially enhancing data security and privacy. As can be seen in **Table 2**, most of the reviewed papers provided similar techniques as primary manners to handle and prevent BD security and privacy issues, including anonymization and encryption. Next useful techniques consist of differential privacy-preserving, cryptography, MapReduce framework, and integrity preservation that are common ways to handle and protect data from malicious attacks and decrease the loss of data during the process of analyzing, interpreting, and synthesizing. Besides, the techniques such as access control, data provenance, multiple computations, fraud detection, auditing, and confidentiality preserving play the third rank in **Table 2**, followed by new recommendations such

as fast anonymization of BD streams, digital watermarking, BD publishing, Cosine protocol, secure storage & sharing, C4.5 algorithm, perturbation, non-synthetic perturbated data generation, and hardware security. Generally, even though those techniques may bring to issues as discussed above, they are also state-of-the-art security and privacy techniques.

Moreover, most of the challenges cover the technical place while researchers do not attract other dimensions. As we can see in **Table 3**, the top ten rows display many researchers agree with the top ten challenges that the Big Data Working Group [11] suggested. This means these ten challenges probably occur and cause unpredictable issues to people, organizations and even governments. It also makes sense of the nature of BD when these challenges involve in technical mechanisms such as distributed programming frameworks, compulsory encryption, etc. In contrast, economics, culture, lack of talent and infrastructure limitation elements may not matter with technicians and persons who conduct technological researches. Therefore, we would like to have some comprehensive implications to enhance the security and privacy of data in the BD scenario below:

1) Although the ethical issue (human intervention) is one of the most important components, which impacts and creates serious problems in the age of BD, this area has not been studied as much as the other ones. Also, every business and technological operation is not isolated with environmental and organizational factors. Therefore, organizations that are dealing with BD should invest in managing and training people how to work ethically and effectively.
2) Like everything, BD has its own life cycle, which includes generating a variety of data, storing, and processing [21]. The techniques reviewed are still fragmentary, and they have not been applied generally to a comprehensive process of BD. Hence, it would like to have a framework of security and privacy for the whole life cycle of BD, which is not only to protect data from external risks but also manage, control, update and maintain data.

Finally, our study obtained all objectives, which include reviewing current issues and preserving techniques, addressing challenges, and giving suggestions. This means the research aims were also achieved completely.

6. Conclusion

It can be concluded that BD plays an important role for not only businesses accomplish their goals by findings of BD analysis, but it also assists governments with appropriate indicators to serve and manage people. However, the current issues have brought many challenges that require a high level of combinations of using technology and other aspects such as ethics, education and training, management to prevent unseen future issues. In this paper, we just reviewed the issues and techniques related to BD security and privacy. At the same time, many other factors affect the quality and quantity of BD as well as require state-of-the-art technologies to protect data. As our future work, evaluating rigorously papers that involve in Big Data needs to be achieved to provide effective solutions of using and protecting data in the BD scenario and especially to the establishment of a new framework for preserving an entire life cycle of BD.

References

[1] I. Foster, R. Ghani, R.S. Jarmin, F. Kreuter and J. Lane, Eds, *Big data and social science: A practical guide to methods and tools*, Boca Raton, FL: CRC Press, 2017.

[2] D. Lazer and J. Radford, “Data ex machina: introduction to big data”, *Annual Review of Sociology*, Vol. 43, pp. 19-39, 2017.

[3] E. Ularu, F. Puican, A. Apostu and M Velicanu, “Perspectives on Big Data and Big Data Analytics”, *Database Systems Journal*, Vol. 3, No. 4, pp. 3-14, 2012.

[4] D. Lazer, R. Kennedy, G. King and A. Vespignani, “The parable of Google Flu: traps in big data analysis”, *Science*, Vol. 343, pp. 1203-1205, 2014.

[5] P. Goes, “Editor’s comments: Big Data and IS Research”, *MIS Quarterly*, Vol. 38, No. 3, pp. 3-8, 2014.

[6] I. Ahmad, T. Haiduwa, A. Kiani and K. Saddique, “Security and Privacy Concerns in Big Data”, *European Academic Research*, Vol. IV, No. 10, pp. 8129-8161, 2017.

[7] B.K. Daniel, “Big Data and data science: A critical review of issues for education

research", *British Journal of Educational Technology*, Vol. 50, No. 1, pp. 101-113, 2019.

[8] J. Gantz and D. Reinsel, *Extracting Value from Chaos*, IDC iView, June 2011. Available: https://www.scribd.com/document/87162672/IDC-IView-Extracting-Value-From-Chaos-2011-Data-Storage-Etc. [Accessed Oct. 7, 2020].

[9] H. Wang, X. Jiang and G. Kambourakis, "Special issue on Security, Privacy and Trust in network-based Big Data", *Information Science*, Vol. 318, pp. 48-50, 2015.

[10] K.C. Laudon and J.P. Laudon, *Management Information Systems: Managing the Digital Firm*, 14th edn, England: Pearson, 2016.

[11] Big Data Working Group, "Expanded Top Ten Big Data Security and Privacy", Cloud Security Alliance (CSA), April 2013. Available: https://downloads.cloudsecurityalliance.org/initiatives/bdwg/Expanded_Top_Ten_Big_Data_Security_and_Privacy_Challenges.pdf. [Accessed Oct. 7, 2020].

[12] H. Leino-Kilpi, M. Valimaki, T. Dassen, M. Gasull, C. Lemonidou, A. Scott and M Arndt, "Privacy: a review of the literature", *International Journal of Nursing Studies*, Vol. 38, pp. 663-671, 2001.

[13] Z. Sun, L. Sun and K. Strang, "Big Data Analytics Services for Enhancing Business Intelligence", *Journal of Computer Information Systems*, Vol. 58, No. 2, pp. 162-169, 2016.

[14] W. Fang, X.Z. Wen, Y. Zheng and M. Zhou, "A survey of Big Data Security and Privacy Preserving", *IETE Technical Review*, Vol. 34, No. 5, pp. 544-560, 2017.

[15] C. Clifton, M. Kantarcioglu, and J. Vaidya, "Defining pri- vacy for data mining," in *Proc. of National Science Foundation Workshop on Next Generation Data Mining*, Baltimore, MD, 2002, pp. 126-133.

[16] Q. Jing, A.V. Vasilakos, J. Wan, J. Lu and D. Qiu, "Security of the Internet of Things: perspectives and challenges", *Wireless Networks*, Vol. 20, No. 8, pp. 2481-2501, 2014.

[17] J.P. Hubaux and A. Juels, "Privacy is dead, long live privacy", *Communications of the ACM*, Vol. 59, No. 6, pp. 39-41, 2016.

[18] K. Kambatla, G. Kollias, V. Kumar and A. Grama, "Trends in big data analytics", *J Parallel Distrib Comput*, Vol. 74, No. 7, pp. 2561-2573, 2014.

[19] G.H. Kim, S. Trimi and J.H. Chung, "Big-data applications in the government sector", *Communications of the ACM*, Vol. 57, No. 3, pp. 78-85, 2014.

[20] H.E. Pence, "What is big data and why is it important?", *Journal of Educational Technology Systems*, Vol. 43, No. 2, pp. 159-171, 2015.

[21] P. Jain, M. Gyanchandani and N. Khare, "Big data privacy: a technological perspective and review", *Big Data*, Vol. 3, No. 25, pp. 1-25, 2016.

[22] J. Bohannon, "Credit card study blows holes in anonymity", *Science*, Vol. 347, No. 6221, pp. 468, 2016.

[23] M.S. Eastin, N.H. Brinson, A. Doorey and G.B. Wilcox, "Living in a big data world: Predicting mobile commerce activity through privacy concerns", *Computers in Human Behaviour*, Vol. 58, pp. 214-220, 2016.

[24] H. Ekbia, M. Mattioli, I. Kouper, G. Avare, A. Ghazinejad, T. Bowman, V.R. Suri, A. Tsou, S. Weingart and C.R. Sugimoto, "Big data, bigger delimmas: A critical review", *Journal of the Association for Information Science and Technology*, Vol. 66, No. 8, 2014.

[25] N. Kshetri, "Big data's impact on privacy, security and customer welfare", *Telecommunications Policy*, Vol. 38, No. 11, pp. 1134-1145, 2014.

[26] J. Jesson, L. Matheson and F.M. Lacey, *Doing your literature review: traditional and systematic techniques*, Los Angeles & London: SAGE Publications, 2011.

[27] H. Arksey and L. O'Malley, "Scoping studies: towards a methodological framework", *International Journal of Social Research Methodology*, Vol. 8, No. 1, pp. 19-32, 2005.

[28] D. Levac, H. Colquhoun and K.K. O'Brien, "Scoping studies: advancing the methodology", *Implementation Science*, Vol. 5, No. 1, pp. 69, 2010.

[29] H.M. Daudt, C. van Mossel and S.J. Scott, "Enhancing the scoping study methodology: a large, inter-professional team's experience with Arksey and O'Malley's framework", *BMC Medical Research Methodology*, Vol. 13, No. 48, pp. 1-9, 2013.

[30] G. Paré, M.C. Trudel, M. Jaana and S. Kitsiou, "Synthesizing information systems knowledge: A typology of literature reviews", *Information and Management*, Vol. 52, No. 2, pp. 183-199, 2015.

[31] M. Templier and G. Paré, "A framework for guiding and evaluating literature reviews", *Communications of the Association for Information Systems*, Vol. 37, No. 6, pp. 112-137, 2015.

[32] H. Cooper and L.V. Hedges, *Research synthesis as a scientific process*, H. Cooper, L.V. Hedges, J.C. Valentine, Eds, 2nd edn, New York: Russell Sage Foundation, pp. 3-17, 2009.

[33] J. Webster and R.T. Watson, "Analyzing the past to prepare for the future: Writing a literature review", *Management Information Systems Quarterly*, Vol. 26, No. 2, pp. xiii-xxiii, 2002.

[34] H.M. Cooper, "Organizing knowledge syntheses: A taxonomy of literature reviews", *Knowledge in Society*, Vol. 1, No. 1, pp. 104-126, 1988.

[35] G. Noblit and R. Hare, *Meta-ethnography: synthesising qualitative studies*, Newbury Park, CA: Sage, 1988.

[36] A. Harden, J. Thomas, M. Cargo, J. Harris, H. Pantoja, K. Flemming, A. Booth, R. Garside, K. Hannes and J. Noyes, "Methods for integrating qualitative

[37] and implementation evidence within intervention effectiveness reviews", in *Cochrane Qualitative and Implementation Methods Group guidance series, J Clin Epidemio*, Vol. 97, pp. 70-78, 2018.

[38] V. Braun and V. Clarke, "Using thematic analysis in psychology", *Quanlitative Research in Psychology*, Vol. 3, pp. 77-101, 2006.

[39] M. Weinhardt, "Ethical Issues in the Use of Big Data for Social Research", *Historical Social Research*, Vol. 45, No. 3, pp. 342-368, 2020.

[40] B. Baesens, R. Bapna, J.R. Marsden, J. Vanthienen and J.L. Zhao, "Transformation Issues of Big Data and Analytics in Network Business", *MIS Quarterly*, Vol. 40, No. 4, pp. 807-818, 2016.

[41] K. Kumari and M. Mrunalini, "A survey on Big Data Security: Issues, Challenges and Techniques", *International Journal of System and Software Engineering*, Vol. 6, No. 2, pp. 23-36, 2018.

[42] J. Moreno, M.A. Serrano and E. Fernández-Medina, "Main Issues in Big Data Security", *Future Internet*, Vol. 8, No. 44, pp. 1-16, 2016.

[43] S. Kim, "Challenges in the management of data in the "Big Data" age", *International Journal of Business Research and Information Technology*, Vol. 2, No. 1, pp. 1-13, 2015.

[44] K.E. Martin, "Ethical issues in the Big Data Industry", *MIS Quarterly Executive*, Vol. 14, No. 2, pp. 67-85, 2015.

[45] P.S. Rani and D. Vigneswari, "Security and privacy in big data analytics", *International Journal on Intelligent Electronic System*, Vol. 10, No. 2, pp. 32-35, 2016.

[46] Z. Sun, K.D. Strang and F. Pambel, "Privacy and security in the big data paradigm", *Journal of Computer Information Systems*, Vol. 60, No. 2, pp. 146-155, 2020.

[47] J. Moura and C. Serrão, "Security and Privacy Issues of Big Data", in *Cloud Security*, 2019, doi: 10.4018/978-1-5225-8176-5.ch080

[48] R. Bao, Z. Chen and M.S. Obaidat, "Challenges and techniques in Big data security and privacy: A review", *Security Privacy*, Vol. 1, No. 13, pp. 1-8, 2018.

[49] K. Shirudkar and D. Motwani, "Securing big data: Security recommendations for hadoop and NoSQL environments", in *Proc. of IEEE*, Vol. 78, pp. 30-56, 2013.

A Suggestion of Deep Learning Clothes Classification System Using Cloud and Edge Computing

Ick Sun Ku[1], Sun-Kuk Noh[2]
1.Daehan eni, Inc., Gwangju, Republic of Korea
2.National Program of Excellence in Software Center, CHOSUN University
Gwangju, Republic of Korea, Corresponding Author
[e-mail: nsk7078@chosun.ac.kr]

Abstract

In Korea, there are many used clothing generation amounts with 61 trillion won of clothing consumption per year, but not only is it not collected properly due to inefficiency of used clothing collection system, but also used clothing collected is not recycled properly due to insufficient recycling system. This paper presents a suggestion of deep learning clothes classification system using cloud and edge computing. The system proposed in this study classifies clothing image datasets input from terminals installed in various places into several classes using convolutional neural network(CNN). And the classification result is stored in Cloud through Edge computing. Edge computing allows data from internet of things devices (terminal) to be analyzed at the edge of the network before being sent to a cloud. This proposed system can be expected to improve the process and automation efficiently in the classification of clothing for recycling.

***Keywords*:** Deep learning, Convolutional neural network, Clothes classification system, Edge computing

1. Introduction

Recently, artificial intelligence (AI), led by machine learning and in-depth learning, has emerged as a key technology of the Fourth Industrial Revolution (4IR). Various research and development related to AI are being carried out in various industrial fields.[1-3]

The textile industry consumes a lot of human resources in all processes such as raw material collection, dyeing, processing, sewing, etc., and the waste of resources and energy and increase of environmental pollution are caused by the short-term waste of clothing produced through this process. For example, 1,500 liters of water is used to make a pair of jeans, and 10 to 15% of the chemicals that come out after weaving and removing dyes become wastewater and pollute the environment. If you buy clothes at Second Hand Shop, which sells used goods instead of buying new clothes, you can reduce environmental pollution as much. Therefore, there are many efforts in domestic and various countries to recycle and re-wear waste clothing if possible. In Korea, there are many used clothing generation amounts with 61 trillion won of clothing consumption per year, but not only is it not collected properly due to inefficiency of used clothing collection system, but also used clothing collected is not recycled properly due to

This research was supported by the MISP (Ministry of Science, ICT & Future Planning), Korea, under the National Program for Excellence in SW (2017-0-00137) supervised by the IITP (Institute of Information & communications Technology Planing & Evaluation)" (2017-0-00137)

insufficient recycling system.[4]

In this study, to make an offer solution about the problem, suggest a deep learning clothes classification system using cloud and edge computing for efficient classification of reuse clothes.

2. Related works

2.1 Reuse clothes

In Korea, clothing waste is not included in the subject of separation collection, and separate discharge and collection are carried out in various forms by local governments.
Waste clothing collected by separation and discharge is classified as reusable and impossible. Reusables are sold in flea markets, bazaars, etc. in Korea and exported to foreign countries including underdeveloped countries.
As shown in **Fig. 1**, 2017 Ministry of Environment statistics show that the waste fibers discharged from the clothing production stage amount to about 224 tons per day, about 8200 tons of waste fibers per year, and non-lung clothing (recyclable clothing) are being exported to Southeast Asia. [5]

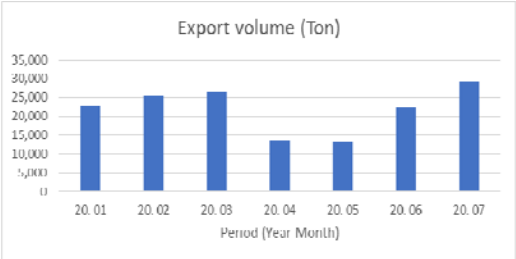


Fig. 1. Reuse clothes export in KOREA

2.2 Clothes classifications with CNN

Artificial intelligence (AI), which is driven by machine learning and deep learning together with new technologies that lead the Fourth Industrial Revolution(4IR), is applied through various research and development. The process of classifying the clothing image dataset using a convolutional neural network (CNN) is shown in **Fig. 2**. In CNN, cloth image datasets were classified into several classes through deep learning. [6-8]

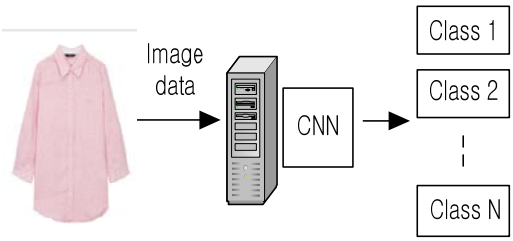


Fig. 2. Clothes classifications with CNN

2.3 Edge computing

Edge computing is a distributed open IT architecture with distributed processing capabilities that enables mobile computing and Internet of Things (IoT) technology. The role of edge computing to date has been mainly used to collect, store, filter and transmit data to cloud systems. Edge computing allows data from internet of things devices (terminal) to be analyzed at the edge of the network before being sent to a cloud. [9]

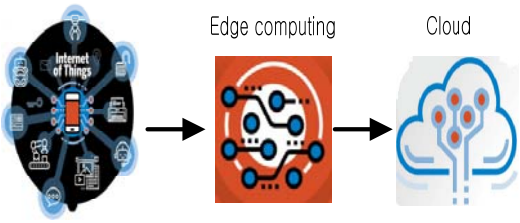


Fig. 3. IoT and Edge computing and Cloud

3. Suggestion of System

The proposed system can be installed in various places where recycled clothing is collected, and the proposed system classifies the clothing image data set input from the terminal into several classes using CNN. Classes can be classified large classes and small classes. Large classes are top and bottom, adult and child, and small classes are such as knit, cardigan, coat, pants, skirt etc. The classification results are transmitted and stored in the cloud through edge computing.

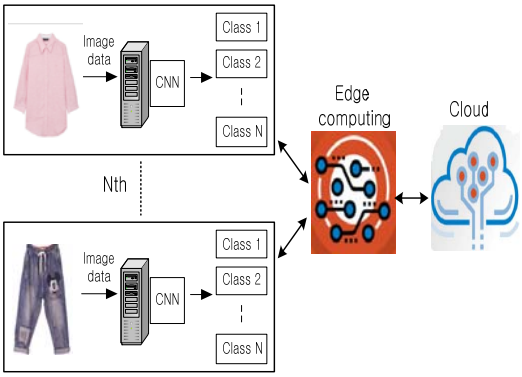


Fig. 4. Suggestion of Clothes Classification with Edge computing and Cloud

4. Conclusions

In Korea, a reuse clothes are not collected properly due to inefficiency of used clothing collection system, but also used clothing collected is not recycled properly due to insufficient recycling system. In this study, suggested a deep learning clothes classification system using cloud and edge computing for efficient classification of reuse clothes. The system proposed classifies clothing image datasets input from terminals installed in various places into several classes using CNN. And the classification result is stored in Cloud through Edge computing. This proposed system can be expected to improve the process and automation efficiently in the classification of clothing for recycling.

Acknowledgment

"This research was supported by the MISP (Ministry of Science, ICT & Future Planning), Korea, under the National Program for Excellence in SW (2017-0-00137) supervised by the IITP (Institute of Information & communications Technology Planing & Evaluation)" (2017-0-00137)

References

[1] Z. Liu, P. Luo, S. Qiu, X. Wang, and X.Tang, "Deepfashion: Powering robust clothes recognition and retrieval with rich annotations," in Proc. IEEE Conf. CVPR, pp.1096-1104, Jun. 2016.

[2] Y. Ge, R. Zhang, X. Wang, X. Tang, and P.Luo, "DeepFashion2: A versatile benchmark for detection, pose estimation, segmentation and re-identification of clothing images," in Proc. IEEE Conf. CVPR, pp. 5337-5345, Jun. 2019.

[3] Itamar Arel, Derek C Rose, Thomas P Karnowski, "Research frontier: deep machine learning a new frontier in artificial intelligence research," IEEE Computational Intelligence Magazine, November 2010.

[4] https://www.gov.kr/portal/gvrnReport/view/H1909000000121708?policyType=G00302&srchTxt=

[5] http://me.go.kr/home/web/board/read.do?pagerOffset=0&maxPageItems=10&maxIndexPages=10&searchKey=title&searchValue=%EC%9D%98%EB%A5%98+%EC%88%98%EC%B6%9C&menuId=286&orgCd=&boardId=1393130&boardMasterId=1&boardCategoryId=&decorator=

[6] Alex Krizhevsky, Ilya Sutskever, GeoffreyE.Hinton,."ImageNet Classification with Deep Convolutional Neural Networks.", COMMUNICATIONS OF THE ACM, VOL. 60, NO. 6, JUNE 2017.

[7] Mohan Laavanya, Veeramani Vijayaraghavan, "Residual Learning of Transfer learned AlexNet for Image Denoising," IEIE Transactions on Smart Processing and Computing, vol. 9, no. 2, April 2020.

[8] Hye-Yeon Son, Min Seon Lee, Sun-Kuk Noh, "A Study on the Clothes Classification using AlexNet Deep Learning," In The 9th International Conference on Smart Media and Applications (SMA 2020), September 2020.

[9] https://www.networkworld.com/article/3224893/what-is-edge-computing-and-how-it-s-changing-the-network.html.

O-Ring Defect Inspection Framework based on Deep Learning Algorithms

Mustafaev Bekhzod[1], Tursunov Anvarjon[2], and JaRok Koo[1]
[1]Department of Electrical and computer Engineering, University of Ulsan (UOU), Ulsan, Korea
[2] Department of Software, Sejong University, Seoul, Korea
[e-mail: bekhzodmustafaev91@gmail.com, tursunovanvarjon@gmail.com, koorok@ulsan.ac.kr]
*Corresponding author: JaRok Koo

Abstract

Abstract: O-ring defect inspection is a challenging task because of their variable sizes as well as defect types. Traditional quality inspection systems use image processing to extract hand-crafted features and machine learning models to classify defect types. However, they are very sensitive to different sizes of O-ring objects and various defect types. Moreover, software maintenance of traditional systems is hard and requires much effort and time. To solve those challenges, we proposed a fully deep learning (DL) – based O-ring defect inspection framework. Proposed system, firstly, uses DL object localization model to find the location of O-ring objects in an image, then inspects it via using DL classification model. We used pre-trained convolutional neural networks (CNN) for both object localization and defect classification and trained using transfer learning technique. The proposed system is evaluated by using our custom O-ring dataset. Object localization was achieved 98.5 % accuracy on test data. 97 % accuracy was achieved on a classification task. Experimental results show that the proposed DL-based system is capable of the inspection of O-ring objects accurately and robust for the various O-ring objects as well as defect types.

Keywords: O-ring defect inspection, Deep Learning, Convolutional Neural Networks (CNNs), Transfer learning

Introduction

O-rings are the small rubber seals that are very widely used in the industry, starting from an individual tiny seal for repairs or maintenance to top-quality assured applications in aerospace, automotive or general engineering. Sealing quality is an important measure for estimating the quality of equipment and directly affects the accuracy and stability of the device. Therefore, vouching for the quality of O-rings is substantially important [1].

Automatic defect inspection using image processing and traditional machine learning is being widely adopted in many industries. The various methods used for these aspects have been reviewed by many studies. Martinez et al. designed a machine vision system for surface quality inspection of transparent parts.[2] Roby et al. [3] compared defect detection algorithms based on Fourier filtering, image convolution, auto-median, and single-step thresholding approaches. Li et al. [4] investigated an appropriate real-time defect detection and location algorithm on surfaces of sequence circular objects. Gaoliang Peng, et al. used series of image processing methods for inspection of O-rings, to improve the localization accuracy of edge detection. [5]

In recent years, Deep Learning-based approaches have yielded very good results for the inspection of various industrial products. In 2012, Krizhevsky et al. made the breakthrough

with AlexNet in object recognition by winning the ImageNet Large Scale Visual Recognition Competition (ILSVRC) first time with CNN.[6] Masci et al. introduced Max-Pooling CNN model approach for supervised steel defect classification [7]. Later, in a different approach, Fatima A. Saiz and Ismail Serrano et al. Combining traditional Machine Learning techniques with CNN has yielded good results and had focused on increasing robustness and accuracy in detecting steel surface defects [8]. Recently, Chao-Ching Ho*, and Eugene Su at. al has been introduced the Deep Learning-based approach of Rubber Gasket Defect Detection using the TensorFlow open-source library [9].

However, traditional image processing and machine learning-based methods for O-ring defect inspection are very sensitive for different sizes of objects and defect types. Moreover, those methods require expert knowledge who has many years of experience in the problem domain. Many experienced inspectors detect O-rings under sufficient lighting conditions, which is low-accuracy, very time consuming and also labor intensity is high [5].

In this paper, we proposed a fully DL-based O-ring defect inspection framework that learns the extraction of useful features along with defect classification jointly. The pretrained CNN models have been used by applying the Transfer learning technique. First, we use object detection to determine the location of the image given by a rectangular, axis-parallel bounding box. Then we crop an object one by one and apply some preprocessing techniques such as gamma correction on images of an O-ring to improve their quality. Finally, the classification part defined to extract, from images, features that can distinguish defective and non-defective O-rings and for classifying them into good or not-good.

The paper is arranged as follows. The Deep learning-based defect inspection method is introduced in Section 2. Section 3 includes experiments and results. In Section 4, conclusions are drawn.

Proposed Method

Traditional image processing-based O-ring defect inspection systems first extract hand-crafted features and then passed to a machine-learning algorithm to detect defects [5]. The main problem of traditional image processing-based defect inspections systems is sensitivity to object sizes as well as defect types. Additionally, the maintenance of these kinds of systems is difficult because of the requirement of much effort and time. We proposed a fully DL-based O-ring defect inspection framework that learns the extraction of useful features along with defect classification jointly. An overview of the proposed O-ring defect inspection framework is illustrated in **Fig. 1**.

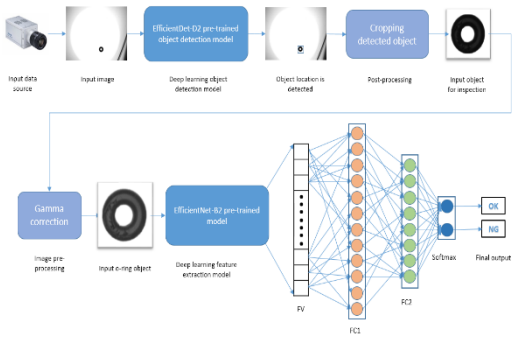


Fig. 1. Proposed fully deep learning (DL) based O-ring defect inspection framework. FV - feature vector, FC – fully connected layer.

To remove redundant data from the input image, firstly, the system detects the O-ring object location in an input image via using DL object detection model. Moreover, this increases the robustness of the proposed system to the sizes of different O-ring objects and locations in an input image. For the detection of the O-ring object in an input image, EfficientDet-D2 [10] pre-trained CNN model was retrained using transfer learning [11] technique. After finding the location of the O-ring object, the post-processing step is applied such as cropping. Input image comes to the system from a high-quality industry camera that is installed on the top of the O-ring inspection machine. The lighting is given from the bottom side of the O-ring rubber. Thus, the O-ring object looks darker than actually, it should be in an input image. In this case, it is difficult to see and detect defects on the surface

of the O-ring objects. To solve this problem, we applied gamma correction [12] in order to fix the lighting condition. As shown in Figure 1, O-ring defect on the surface can be seen clearly after applying gamma correction. We can control brightness of an image with gamma value. If gamma value is higher than 0.5 then an image will be lighter. If gamma value is less than 0.5 then the processed image will be darker than original image. The formula for calculating the resulting output of gamma correction is given Equation 1. Where I is input image, γ is gamma value to change the brightness, and I' is output image.

$$I' = 255 \times \left(\frac{I}{255}\right)^{\gamma} \quad (1)$$

The next step is extraction of robust salient features from O-ring input image by using EfficientNet-B0 [13] pre-trained CNN model and classifying it into "*OK*" and "*NG*" classes. Instead of final fully connected layer of EfficientNet-B0 pre-trained CNN model, we used 2 fully connected layer with 128 and 64 hidden nodes respectively. We used transfer learning [11] technique to train DL classification model with modified fully connected layers. The reason to use pre-trained CNN models and transfer learning is that pre-trained CNN models are already optimized with big dataset to solve general image classification problem. Our dataset was not big thus using pre-trained models will provide high performance even with small number of samples.

Experiments and Results.

3.1 Data collection and preparation

If we want to build robust defect inspection system using deep learning (DL) algorithms, we have to fill several DL requirements. For instance, number of training samples should be sufficient in each class for DL model to learn and generalize the solution to problem. There are many public image datasets such as ImageNet [14] for computer vision problems such as object detection, image classification, and object localization. However, when it comes to solve computer vision problems such as object localization and defect inspection in industry, we have to build our own dataset. Dataset was collected via taking pictures from industrial camera which was installed in factory. There are 40 different sizes of O-ring objects. As DL models require more samples for better accuracy, number of defect samples in each size of O-ring object should be sufficient. To fulfill this requirement, we used 2 types of image augmentation method to increase number of samples in dataset. First one is zooming. In this method we zoomed small sizes of O-ring objects 1.5 and 2 times. The second method is creating artificial defect samples using photoshop program.

Table 1. Detailed description of O-ring objects dataset. Good O-ring objects given as OK and not good O-ring objects as NG.

	OK	**NG**
Train	1071	2439
Validation	120	272
Test	133	286
Total	**1324**	**2997**

The second method was applied when number of defect samples were very few in certain size and defect types of O-ring objects. Finally, we have enough number of O-ring samples for training DL object localization and classification models. Detailed description of dataset is given in **Table 1**.

The next step is labelling O-ring objects after data collection. O-ring objects without defects were labeled as *OK* and O-ring objects which have defects were labeled as *NG*. We used labelImg [15] for labeling all O-ring objects for object localization.

3.2 Experimental setup

For DL models training, we used TensorFlow (version 2.2.0) DL framework and Python (with version 3.7.0) programming language. First of all, we trained DL model for object localization. We used EfficientDet-D2 [10] pre-trained convolutional neural network (CNN) for transfer learning. As object detection model does both object localization and classification, we can use only object detection model for defect inspection. However, O-ring defects are not easy to inspect, so object detection model was used only for object

localization. Training parameters for object detection model were set as in the following: batch size was 10, learning rate was 0.0005, and momentum was 0.99. Model was trained on NVIDIA GeForce RTX 2060 GPU which has 6 GB of graphic memory.

EfficientNet-B0 [13] pre-trained CNN model was retrained by changing its last fully connected layer with 2 custom fully connected layers. We applied transfer learning [11] technique to train pre-trained CNN models. Because those models are already trained with big data to solve image classification problem. They are very good for solving general image classification and object detection problems. Additionally, we do not have large dataset to train new CNN model and achieve high accuracy. EfficientNet-B0 [13] pre-trained CNN model gave high classification accuracy after retraining with a few epochs. Training parameters were set as following: batch size was 64, epochs were 50, activation function was RELU, and optimizer was stochastic gradient descent (SGD). The size of feature vector (FV) is 1280, first fully connected layer (FC1) has 128 hidden nodes, and the second fully connected layer (FC2) has 64 hidden nodes. For training 80 % of overall data was used, other 10% data for validation, and the rest 10% was used for testing in both object detection and classification models training.

3.3 Experimental results

Fig. 2 shows the result of object detection model on test data. It can be seen from the **Fig. 3** that localization error is only 1.5 %. As we used DL object detection only for object localization, 98.5 % localization accuracy is quite sufficient for applying this model for finding O-ring objects location in an image in real time.

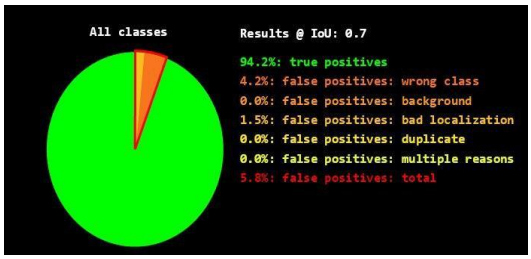


Fig. 2. Object detection result on test data.

Training and validation accuracy result is given in **Fig. 3**. **Fig. 3** illustrates that DL classification model learned well during training without overfitting and underfitting problems and achieved over 95 % of accuracy for both training and validation data. Training and validation loss result is shown in Figure 4. Both training and validation losses were decreased gradually during optimization of DL classification model. Model gave good classification accuracy on both training and validation data. When DL model gives accurate result on unseen data, then it can be applied to solve real world problem. Testing result of DL classification model is given in **Table 3**. It can be seen from **Table 3** that DL classification model gave 97 % of accuracy on test data. This means that DL classification model can be applied to solve real world O-ring defect inspection problem.

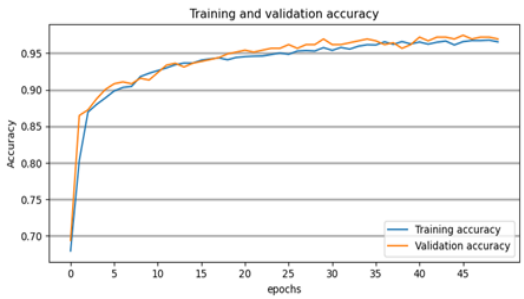


Fig. 3. Training and Validation accuracy result during the training.

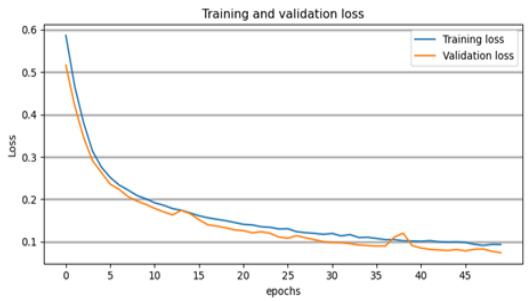


Fig. 4. Training and Validation loss result during the training.

Table 2. The confusion matrix of the original labels versus our predicted labels.

Confusion Matrix

Actual	Classes	NG	OK
	NG	281	5
	OK	7	126

Predected

Table 3. Testing result of DL classification model.

Classes	**Precision**	**Recall**	**F1**	**Num of images**
NG	98%	98%	98%	286
OK	96%	95%	95%	133
Total Accuracy			**97%**	**419**

Conclusion

In this work, O-ring defect inspection system was built using 2 DL models. Object detection model achieved 98.5 % accuracy on object localization problem. We used transfer learning [11] technique for both object localization and O-ring defect inspection problems. We used EfficientDet-D2 [10] pre-trained CNN model for object localization. For classification problem we used EfficientNet-B0 [13] pre-trained CNN model. After training EfficientNet-B0 [13] CNN model, it achieved 97 % accuracy on test data.

References

[1] Trelleborg Sealing Solutions "O-Rings" Available:https://www.tss.trelleborg.com/en/products-and-solutions/o-rings, Edition September, 2020.

[2] Martínez, S.S., Ortega, J.G., García, J.G., García, A.S. and Estévez, E.E., An industrial vision system for surface quality inspection of transparent parts. *The International Journal of Advanced Manufacturing Technology*, *68*(5-8), pp.1123-1136, 2013.

[3] Boby, R.A., Sonakar, P.S., Singaperumal, M. and Ramamoorthy, B.,. Identification of defects on highly reflective ring components and analysis using machine vision. *The International Journal of Advanced Manufacturing Technology*, *52*(1-4), pp.217-233, 2011.

[4] Li, Lin, Zhong Wang, and Fang-ying Pei. "Vision-based surface defect inspection for sequence circular objects." In *2012 5th International Congress on Image and Signal Processing*, pp. 1352-1356. IEEE, 2012.

[5] Peng, G., Zhang, Z. and Li, W, Computer vision algorithm for measurement and inspection of O-rings. *Measurement*, *94*, pp.828-836, 2016.

[6] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." In *Advances in neural information processing systems*, pp. 1097-1105, 2012.

[7] Masci, Jonathan, Ueli Meier, Dan Ciresan, Jürgen Schmidhuber, and Gabriel Fricout. "Steel defect classification with max-pooling convolutional neural networks." In *The 2012 International Joint Conference on Neural Networks (IJCNN)*, pp. 1-6. IEEE, 2012.

[8] Saiz, Fátima A., Ismael Serrano, Iñigo Barandiarán, and Jairo R. Sánchez. "A Robust and Fast Deep Learning-Based Method for Defect Classification in Steel Surfaces." In *2018 International Conference on Intelligent Systems (IS)*, pp. 455-460. IEEE, 2018.

[9] Chao-Ching, H., Su, E., Li, P.C., Bolger, M.J. and Pan, H.N, Machine Vision and Deep Learning Based Rubber Gasket Defect Detection. *Advances in Technology Innovation*, *5*(2), p.76, 2020.

[10] Tan, Mingxing, Ruoming Pang, and Quoc V. Le. "Efficientdet: Scalable and efficient object detection." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10781-10790. 2020.

[11] Hussain, Mahbub, Jordan J. Bird, and Diego R. Faria. "A study on CNN transfer learning for image classification." In *UK*

Workshop on Computational Intelligence, pp. 191-202. Springer, Cham, 2018.

[12] Rahman, S., Rahman, M.M., Abdullah-Al-Wadud, M., Al-Quaderi, G.D. and Shoyaib, M, An adaptive gamma correction for image enhancement. *EURASIP Journal on Image and Video Processing*, pp.1-13, *(1) 2016.*

[13] Tan, Mingxing, and Quoc V. Le. "Efficientnet: Rethinking model scaling for convolutional neural networks." *arXiv preprint arXiv:1905.11946* (2019).

[14] Deng, Jia, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. "Imagenet: A large-scale hierarchical image database." In *2009 IEEE conference on computer vision and pattern recognition*, pp. 248-255. Ieee, 2009.

[15] Darrenl tzutalin, “labelImg” Available: https://github.com/tzutalin/labelImg

Prediction of ripeness stages of fruit trees by combining machine learning techniques and deep learning networks

Sangkyoon Kim[1], Wanhyun Cho[2*], Myung Hwan Na[2], and Inseop Na[3]
[1]Department of Electronic Engineering, Mokpo National University, Muan, Jeonnam, South Korea
[2]Department of Statistics, Chonnam National University, Gwang-Ju, South Korea
[3] National Program of Excellence in Software Centre, Chosun University, Gwang-Ju,South Korea
[e-mail: narciss76@mokpo,ac,kr, {whcho, nmh}@chonnam.ac.kr, ypencil@hanmail.net]
*Corresponding author: Wanhyun Cho

Abstract

In this paper, we propose a prediction system that can automatically determine the ripening stage of fruit trees by combining the pattern recognition techniques and the deep learning networks. First, we use various CNN algorithms to extract color feature vectors most suitable for classifying them from the observed images representing each ripening stage of fruit trees. Here, we used ImageNet, VGG Net, Inception v3, ResNet and Mobilenet as CNN algorithms to extract feature vectors. Second, we applied various pattern classification methods to determine the ripening stage of fruits based on the extracted color features. Here, we have considered SVM, KSVM, K-Nearest Neighborhood, Decision Tree, Random Forest, Gradient Boosting Regression Tree, and Multilayer Perceptron. Finally, we propose a prediction system that can automatically and accurately determined their maturity through experimental images of various fruits such as strawberries and tomatoes by combining CNN and pattern recognition methods. From the experimental results, we can first see that the combination of VGG Net and pattern recognition method showed a high classification rate. Second, we can also see that the combination of the Mobilenet model and the pattern recognition method showed a good classification rate.

Keywords: Harvest prediction system, Fruit ripeness image, Convolution Neural network, Patter recognition methods, Tomato and strawberry image.

1. Introduction

Among various fruits, strawberries or tomatoes are one of Korean's favorite fruits and are used as additives in various foods. Therefore, consumption of these fruits is increasing year by year, making them the most economically value-added fruits. For this reason, strawberries and tomatoes have become the most popular fruits for farmers to grow.

Recently, farmers who grow strawberries and tomatoes are interested in how to grow smart farms as a way to improve the quality and added value of fruit trees. Smart farm technology is a cultivation technology that automates the entire process from fruiting to sowing, growth and harvesting. We think the most important factor in these smart farm cultivation technologies is the automatic determination of the exact timing of harvesting of fruit trees. Therefore, there is a need to develop a prediction system that can automatically determine the harvest time.

This work was partially supported by the Research Program of Rural Development Administration (Project No. PJ015361012020), and the Korea National Research Foundation (Project No. 2020R1F1A1067066).

Here, a large number of training images are needed to use a deep learning network such as CNN to predict the exact harvest time of fruits such as strawberries and tomatoes. However, it is practically impossible for each farmer to collect a large number of images necessary for sorting fruits by each ripening stage at the cultivation field. Therefore, we have to think of deep learning networks that either generate more images from properly collected images or do not need so many training images. For this reason, in this paper, we considered a new deep learning network that can automatically predict the harvest time of fruits by combining deep learning networks such as CNN and machine learning algorithms used in existing pattern recognition. In Section 2 of this paper, we describe the method of collecting fruit image data necessary for our research, propose a structure of a new classification system used to classify them, and suggest a method to learn them. And in Section 3, the experimental methods and results for the performance evaluation of the proposed system are presented, and the performance of the existing prediction systems is compared and analyzed. Finally, in Section 4, we comprehensively analyze and summarize the results obtained through this study to clearly present the results of this study.

2. Dataset and Methods

In this section, we introduce a method of collecting fruit image data to be used for analysis. And the deep learning network used to extract feature vectors from the pretrained CNN model is briefly described, and classification methods to be used to classify the ripeness of fruits using the extracted feature vectors are introduced. In addition, the structure of the predictive system for harvesting time proposed in this paper is presented and the learning method of the proposed system is described.

2.1 Dataset

First, in order to evaluate the performance of the proposed prediction system, we downloaded various strawberry images that were published on the Internet. **Fig. 1** below shows examples of the strawberry images of the 4-stages ripeness level. They are unripe, partially ripe, ripe and over ripe. Moreover, we collected approximately 60 images for each ripeness stage and constructed a total of 250 image data.

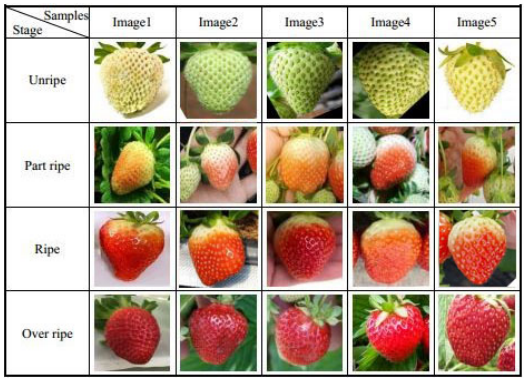


Fig. 1. Examples of image data for each ripening stage of strawberries.

Second, we also downloaded various tomato images that were published on the Internet and used them as experimental data. **Fig. 2** below shows examples of the tomato images of the 4-stages maturity level. Moreover, in order to evaluate the performance of the proposed methods, we collected approximately 75 images for each ripeness stage and constructed a total of 300 image data.

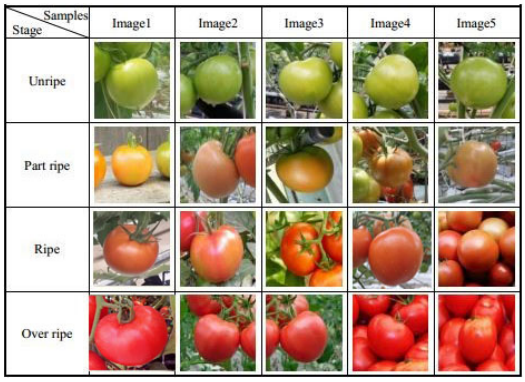


Fig. 2. Examples of image data for each ripening stage of tomato

2.2. Proposed prediction system

Here, we present the Convolution Neural Network and Pattern Recognition Method (CNN-PRM) combined model for prediction the ripeness of fruit tree. The framework for proposed sytem is shown in **Fig. 4**.

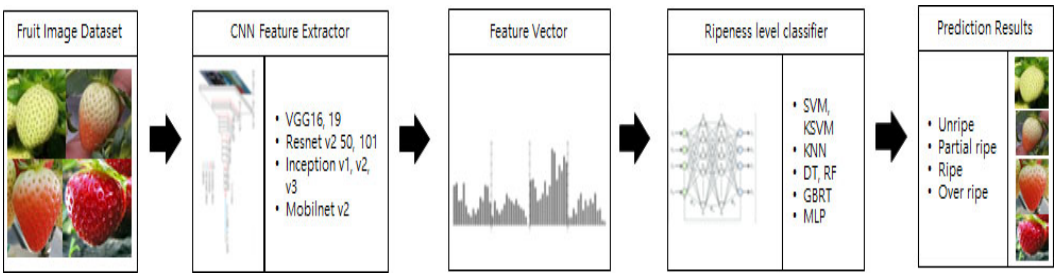


Fig. 4. Architeture of the proposed CNN-PRM predition system

2.2.1. CNN-based feature extraction

Here, various pre-trained CNN models with alternating convolutional and max-pooling layers are used to extract deep features. At this time, the deep picture extracted from the CNN model is used as an input to the pattern recognition methods used for classification. At this time, as the CNN model used to extract the deep picture, a total of 8 models such as VGG Net, ResNet, Inception Net, and MobileNet considered above are used.

2.2.2. PRM-based ripeness prediction

Here, we remove the last output layer of pre-trained CNN models and replace it with a pattern recognition method to be used for classification. At this time, the output values of the last node of the CNN models from which the last layer has been removed are used as input values for pattern recognition methods for ripeness prediction. The pattern recognition methods we will use for prediction are a total of 7 methods, such as SVM, KSVM, K-nearest neiberhood, Decision Tree, Random Forest, Gradient Boosting Regression Tree, Multilayer Perceptron discussed on above.

3. Experimental results

Here, in order to evaluate the performance of the proposed prediction model, we conducted an experiment using strawberry and tomato image data obtained from the Internet.

3.1 Strawberry images data

Here, in order to train the proposed classification methods, we have used 30 images as training dataset for each ripening stage. We also used 20 images as testing data for each stage to verify the performance of trainned prediction system. The following **Table 1** shows the results of evaluating the classification performance for each of combinations of these CNN models and pattern recognition methods for strawberry. First, from the results of **Table 1**, the model combining vgg16 and mlp showed the highest classification rate at about 90%, and the model combining Inception v2 and ksvm showed the second highest classification rate at about 85%. Second, the combined model of Vgg19 and pattern recognition methods such as knn, RF, and BGT showed a higher overall recognition rate than other methods.

Table 1. Correct prediction rate for stawberry image

	svm	ksvm	knn	DT	RF	GBT	MLP
Vgg16	**0.84**	0.80	0.69	0.68	0.77	0.73	**0.90**
Vgg19	0.77	0.81	**0.70**	0.70	**0.79**	**0.79**	0.83
Resnet 50	0.80	0.71	0.61	**0.75**	0.69	0.59	0.79
Resnet 101	0.72	0.74	0.59	0.59	0.68	0.66	0.78
Inception v1	0.79	0.77	0.60	0.52	0.68	0.67	0.79
Inception v2	0.82	**0.85**	0.69	0.53	0.73	0.62	0.84
Inception v3	0.77	0.78	0.62	0.58	0.65	0.56	0.73
Mobilenet v2	0.80	0.78	0.69	0.60	0.68	0.70	0.84

3.2 Tomato images data

Here, in order to train the proposed classification methods, we have used 40 images for each ripening stage as learning dataset. We also used 20 images as the testing data for each stage to verify the perormance of trained classification methods. The following **Table 2** shows the results of evaluating the classification performance for each of combinations of these CNN models and pattern recognition methods for tomato. First, from the results of Table 2, The model combining vgg19 and svm showed the

highest classification rate at about 84%, and the model combining vgg19 and ksvm and MLP showed the second highest classification rate at about 80%. Secondly, the unique thing is that the combined model of mobilenet v2 and several pattern recognition methods generally showed a high recognition rate.

Table 2. Correct prediction rate for tomato image

	svm	ksvm	knn	DT	RF	GBT	MLP
Vgg16	0.73	0.73	0.58	0.64	**0.72**	0.60	0.78
Vgg19	**0.84**	**0.80**	0.65	0.60	0.70	0.59	**0.80**
Resnet 50	0.58	0.63	0.48	0.47	0.46	0.47	0.60
Resnet 101	0.77	0.75	0.53	0.55	0.63	**0.69**	0.73
Inception v1	0.72	0.74	0.60	0.52	0.60	0.61	0.77
Inception v2	0.68	0.68	0.48	0.48	0.61	0.43	0.69
Inception v3	0.74	0.68	0.58	0.44	0.63	0.52	0.70
Mobilenet v2	0.76	0.76	**0.67**	**0.67**	0.61	0.62	0.77

When the experimental results on the classification of strawberries and tomatoes by maturation stage are summarized, it can be seen that the combination of vggnet and pattern recognition method shows a high classification rate. And when developing an app for smartphones that can be used easily in the field, combining the mobilenet model and the pattern recognition method is expected to improve the classification rate, as well as the farmers can easily use it to determine the maturation stage.

4. Conclusions

In this paper, we proposed a prediction system that can automatically discriminate the ripening stages of fruits such as strawberries and tomatoes. The proposed prediction system is constructed by combining various CNN models and pattern recognition methods. In this system, the CNN model was used to extract feature vectors for classification from fruit tree images, and the pattern recognition method was used as a method of classifying the ripening stages of fruit trees using the extracted feature vectors.

The experiment was performed based on the strawberry images and tomato images we collected, and the following results were obtained. First, it was found that the combination of VGGnet and pattern recognition method showed a high classification rate. Second, the combination of the Mobilenet model and the pattern recognition method showed a good classification rate.

In the future, we intend to apply the proposed algorithm to various fruit trees and also develop a smartphone app that farmers can easily use in the field.

References

[1] E. Mavridou, E. Vrochidou, G. A. Papakostas, T. Pachidis and V. G. Kaburlasos, “Machine Vision Systems in Precision Agriculture for Crop Farming”, *Journal of Imaging*, Vol 5, No. 89, pp 1-32, 2019.

[2] R. Pandy, S. Naik, and R. Marfatia, “Image Processing and Machine Learning for Automated Fruit Grading System: A Technical Review,” *International Journal of Computer Application*, Vol. 81. No. 16, pp. 29-39, 2013.

[3] M. De-la-Torre, O. Zatarain, H. Avila-George, M. Munoz, J. Oblitas, R. Lozada, J. Mejia and W. Castro, “Multivariate Analysis and Machine Learning for Ripeness Classification of Cape Gooseberry Fruits”, Processes, Vol 7, N0. 928, pp 1-15, 2019.

[4] V. Nambi, K. Thangavel, A. Manickavasagan, and S. Shahir, “Comprehensive ripeness-index for prediction of ripening level in mangoes by multivariate modeling of ripening behavior”, *International Agrophysics*, Vol. 31, pp 35-44, 2017.

[5] J. Naranjo-Torres, M. Mora, R. Hernandez-Garcia, R. J. Barrientos, C. Fredes and A. Valenzuela, “A Review of Convolution Neural Network Applied to Fruit Image Processing, *Applied Sciences*, Vol. 10, No. 3443, pp 1-31, 2020.

[6] Z. Gao, Y. Shao, G. Xuan, Y. Wang, Y. Liu and X. Han, “Real-time hyperspectral imaging for the in-field estimation of strawberry ripeness with deep learning”, *Artificial Intelligence in Agriculture*, Vol. 4, pp 31-38, 2020.

KSII The 12th International Conference on Internet (ICONI) 2020.

Two Types of Sharing Task View in Remote Collaboration Study

[1]**Seungwon Kim and [2]In Seop Na**
[1]STEM, University of South Australia, Adelaide, Australia
[2] National Program of Excellence in Software Centre, Chosun University, Gwangju, South Korea
[e-mail: Seungwon.Kim@unisa.edu.au, ypencil@hanmail.net]
Corresponding author: In Seop Na

Abstract

Remoter collaboration has been studied for more than two decades. One main topic in remote collaboration studies is providing a tool of sharing task view to a remote expert from local end system. In this paper, we describe previous studies on sharing task view. There are mainly two types of sharing task view: dependent and independent views. With the dependent view, a remote expert and local worker have an identical view controlled by a local worker. With the independent view, both local worker and remote expert have individual view controlled by each of them.

***Keywords*:** Augmented Realty, Remote Collaboration, Sharing Task View

1. Introduction

Collaboration is multi-person activities to achieve a common goal [1]. If collaborators are distributed at different places or environments, it is remote collaboration.

Many researchers reports that remote collaboration is more difficult than co-located collaboration because the communication and understanding task in remote collaboration is limited compared to the one in co-located collaboration [2,3]. Therefore, researchers in remote collaboration study have researched on how collaborators have better communication and better understanding the task [4,5].

Since a remote user is not in the task space, there should be a way showing current state of the task [2] to the remote user. In this paper, we discuss previous studies exploring the effect of sharing task view by categorizing them into two types: dependent and independent views.

2. Views for Remote Collaboration

In remote collaboration, the system supports sharing the local task view between collaborators. When sharing the local space, a remote expert can view the local space and the local worker's task activities [2] and also use visual communication cues to communicate back.

2.1 Dependent View

For providing dependent view in remote collaboration systems, researchers used either a head mounted camera (HMC) or a tablet and allowed a local worker to control the viewpoint of a live video that is shared with a remote expert [6,7]. A local worker wore a HMC on his or her head [7,8] or held a tablet for using its camera [9]. With the HMC or tablet, both local worker and remote expert had an identical view, and the view was controlled by local worker's head or hand movement.

The dependent view has a benefit of WYSIWIS (What You See Is What I See) [10]. Since collaborators have an identical view, they do not worry about the collaborating partner watching different area of the task space and missing any shared activities such as a local worker manipulating an object, or a remote expert using and showing visual cues. This increases the level of awareness of the shared activities and collaborators do not need to spend any effort for aligning their viewpoints during the collaboration [6].

However, these benefits of the dependent view have trade-offs [1,2]. During the collaboration, collaborators may need to conduct individual activities (i.e. looking at different area of a task space or the local worker posing in a comfortable pose during collaboration without worrying about remote expert's view) but the dependent view limits these individual activities. Moreover, sudden viewpoint changes with the dependent view can raise an issue of motion sickness to the remote expert [2].

2.2 Independent View

To solve the issue of the dependent view and provide freedom to look around the task space, researchers employed the independent view. With the independent view, a remote expert can have viewpoints different from the local user and separately navigate around the scene. Some researchers developed a hardware system in which a remote expert controls a camera in the local environment [11,12].

Some others implemented software functions providing an independent view with the source images of the live video from local worker's HMC or tablet [6,9]. A simple method used was saving a still image from the live video at an important viewpoint or moment and retrieving it when a remote expert needs to have an independent view [8,9].

The method of using still images from the live video has been advanced with additional functions. First, some researchers automatically paused the live video when a remote expert starts using the visual cues (called auto-freeze function) [6], so the remote expert could use visual cues within the paused still image view without worrying about sudden viewpoint change. Second, some researchers stitched the images from a live video and then constructed a large image covering the overall task space [13,14]. When a local worker is looking around the task space and a new frame image contains a new area that was not stitched into the large image, the system stitched it into the large image. With the large image, a remote expert could have an independent view regardless of the current viewpoint of local worker's live video and navigate around the task space. However, the stitching image and constructing a large image view has a limitation: a remote expert will not be able to see temporal changes at the area where the new frame is not stitched on.

To solve the limitation, some researchers used a 360-degree panoramic camera to provide a real-time 360 independent view and allowed a remote expert to freely look around [15-18]. However, if the 360 camera is physically mounted on the local worker's head [19] or shoulder [16], the 360 live video is rotated depending on the local worker's head and shoulder rotation. To solve the issue, Lee et al. [17] used a tracking sensor and counter-rotated the 360 live video in opposite direction to the local worker's rotation before the system rendered the 360 live video scene. In this way, the remote expert view remained at a fixed orientation and only reflected the remote expert's view rotation. Another issue with the 360 independent view is difficulty in understanding each other's view direction. While both local and remote collaborators can independently look around the shared scene, it becomes difficult for the collaborators to tell where the other is looking at. To overcome this difficulty, Lee et al. [17] implemented view awareness cues such as view frame or arrows pointing at where the other user is looking at. Despite of these benefits, the shared live 360 video is a 2D presentation that has limited depth perception and it can only provide views [16-18] independent only in orientation but not in position.

Recently, researchers starts reconstructing the task space and share it so the remote expert can have a 3D view which is almost identical with the local task space. The 3D view supports depth perception and allows to have the independent view at different position and orientation [19].

Some researchers explored effect of 3D view in remote collaboration study [19]. In remote collaboration studies, researchers studied the effect of 3D reconstruction view during the collaboration [15], visual tools increasing awareness where the collaborating partner is looking at [19], and the method of representing the remote expert (such as avatar) in local environment [15].

3. Conclusions

From the review, we found that recent studies mostly used HMDs on both local worker and remote expert ends. Researchers developed Mixed Reality systems that share 3D reconstructed local environment where a remote expert collaborates in the identical virtual environment with the one that a local worker has in the real world.

References

[1] C. Gutwin, and S. Greenberg, "Design for individuals, design for groups: tradeoffs between power and workspace awareness" in *Proc. of ACM conference on Computer supported cooperative work*, pp. 207–216, 1998.

[2] S. Kim, M. Billinghurst and G. Lee "The effect of collaboration styles and view

independence on video-mediated remote collaboration", *Computer supported cooperative work,* vol. 27, no 3-6, Dec. 2018. pp 569-607.

[3] W. Huang, S. Kim, M. Billinghurst and L. Alem, "Sharing hand gesture and sketch cues in remote collaboration" *Journal of Visual Communication and Image Representation,* vol. 58, Jan. 2019, pp 428-438.

[4] S. R. Fussell, L. D. Setlock, J. Yang, J. Ou, E. Mauer, and A. D Kramer, "Gestures over video streams to support remote collaboration on physical tasks" *Human-Computer Interaction*, vol. 19, no. 3, pp. 273-309. November 2009

[5] S. Kim, M. Billinghurst, C. Lee, and G. Lee, G. "Using Freeze Frame and Visual Notifications in an Annotation Drawing Interface for Remote Collaboration". *KSII Transactions on Internet & Information Systems*, Vol. 12 no. 12, 2018

[6] S. Kim, G. A. Lee, S. Ha, N. Sakata, and M. Billinghurst, "Automatically freezing live video for annotation during remote collaboration" in *Proc. of ACM CHI Extended Abstracts*, pp. 1669-1674, 2015

[7] L. Gao, H. Bai, T. Piumsomboon, G. Lee, R. W. Lindeman, and M. Billinghurst, "Real-time visual representations for mixed reality remote collaboration" *In Proc. of International Conference on Artificial Reality and Telexistence & Eurographics Symposium on Virtual Environments, 2017*

[8] L. Alem, F. Tecchia and W. Huang, W. "Remote tele-assistance system for maintenance operators in mines", *In Proc. of 11th Underground Coal Operators' Conference*, pp 171 -177, 2011

[9] S. Kim, G. A. Lee, N. Sakata, and M. Billinghurst, "Comparing pointing and drawing for remote collaboration", *In Proc. of IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 1-6). 2013

[10] M. Stefik, D. G. Bobrow, G. Foster, S. Lanning, and D. Tatar, "WYSIWIS revised: early experiences with multiuser interfaces" *ACM Transactions on Information Systems (TOIS)*, Vol. 5, No. 2, April 1987, pp. 147–167

[11] H. Kuzuoka, G. Ishimoda, Y. Nishimura, R. Suzuki, and K. Kondo "Can the GestureCam be a Surrogate?" *In Proc. of the Fourth European Conference on Computer-Supported Cooperative Work ECSCW'95* (pp. 181-196). 1995

[12] N. Sakata, T. Kurata, T. Kato, M. Kourogi, and H. Kuzuoka, "WACL: Supporting Telecommunications Using Wearable Active Camera with Laser Pointer" *In Proc. of Seventh IEEE International Symposium Wearable Computer*, pp 53-56, 2003

[13] S. Kasahara, and J. Rekimoto, "JackIn: integrating first-person view with out-of-body vision generation for human-human augmentation" *In Proc. of the 5th Augmented Human International Conference* 2014

[14] S. Gauglitz, B. Nuernberger, M. Turk, and T. Höllerer, "World-stabilized annotations and virtual scene navigation for remote collaboration" *In Proc. of the 27th annual ACM symposium on User interface software and technology*, pp. 449-459, 2014

[15] T. Teo, L. Lawrence, G. A. Lee, M. Billinghurst, and M. Adcock "Mixed Reality Remote Collaboration Combining 360 Video and 3D Reconstruction" *In Proc. of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019

[16] T. Piumsomboon, G. A. Lee, A. Irlitti, B. Ens, B. H. Thomas, and M. Billinghurst, "On the Shoulder of the Giant: A Multi-Scale Mixed Reality Collaboration with 360 Video Sharing and Tangible Interaction" *In Proc. of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019

[17] G. A. Lee, T. Teo, S. Kim, and M. Billinghurst, "A user study on MR remote collaboration using live 360 video", *In Proc. of IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 153-164, 2018,

[18] A. Tang, O. Fakourfar, C. Neustaedter, and S. Bateman, "Collaboration in 360° videochat: Challenges and opportunities" *In Proc. of the 2017 Conference on Designing Interactive Systems, 2017*

[19] T. Piumsomboon, A. Dey, B. Ens, G. Lee, and M. Billinghurst, "The effects of sharing awareness cues in collaborative mixed reality", *Frontiers in Robotics and AI*, vol 6, no 5, Feb. 2019, pp 5

RECP from the perspective of Chinese and Taiwanese Netizens : a comparative Semantic Network analysis

Shaopeng Che, Pim Kamphuis, and Jang Hyun Kim
Department of Interaction Science/Department of Human-Artificial Intelligence Interaction,
Sungkyunkwan University
Seoul, South Korea
[e-mail: cheshao@g.skku.edu, p.i.m@skku.edu]
*Corresponding author: Jang Hyun Kim

Abstract

In November 2020, fifteen countries signed the Regional Comprehensive Economic Partnership (RCEP), shaping a new free-trade zone in Asia and the Pacific by joining together some of the largest economies in the world such as South Korea, Japan, and China. Despite being geographically located in the middle of this trade-zone, Taiwan is excluded from this agreement. To understand how it perceives the exclusion and explore the concerns of the Taiwanese people, we have crawled comments from *Sina Weibo* in China and *YouTube* in Taiwan that contained the keyword "RECP". Semantic Network Analysis shows that Chinese netizen's comments mainly discuss the economic impact, while Taiwanese netizens seem to be mostly concerned about China.

***Keywords*:** RCEP, China, Taiwan, SNA, differences, clustering analysis

1. Introduction

On November 15, 2020, fifteen participating countries concluded the Fourth Regional Comprehensive Economic Partnership Agreement Leaders' Meeting by signing an agreement that cements the world's largest plurilateral trade agreement, which accounts for more than 30% of the worldwide population [1]. Eliminating a significant amount of tariffs on import and export goods among the countries, the signing of the RCEP is an important step for regional countries in safeguarding their multilateral trading system, building up to an open world economy. It is of symbolic significance in deepening regional economic integrations and stabilization of the global economy.

Fig. 1. RCEP coverage
Source: https://www.google.com/url?sa=i&url=https%3A%2F%2Fasiatimes.com%2F2020%2F11%2Frcep-trade-pact-heralds-dawn-of-asian-century%2F&psig=AOvVaw1r94lvEMz-NvxtGC30KjZq&ust=1606895436368000&source=images&cd=vfe&ved=0CAIQjRxqFwoTCLjUoZSmrO0CFQAAAAAdAAAAABAE

Interestingly, the new economic trade zone shaped by the RCEP does not include Taiwan,

which is located right in the middle of the new trading bloc. Taiwanese netizens tend to laugh and jokingly exclaim that RCEP is an economic circle build around Taiwan - without Taiwan.

Based on this fact that the establishment of RCEP has just been completed with Taiwan in isolation, this paper attempts to determine the differences in perception between Chinese and Taiwanese netizens, and to understand how citizens of countries that are excluded from such agreements could deal with this in the future.

2. Method

2.1 Data crawling

- Data sources: we used web scrapers to collect comments from *Sina Weibo* in China and *YouTube* in Taiwan [2]
- Search keywords: only comments that mentioned “RCEP” in China and "RCEP 臺灣" in Taiwan were scraped
- Data range: comments posted from November 15, 2020 to November 17, 2020
- Data size: 5598 comments from China and 5751 comments from Taiwan

2.2 Data preprocessing

- Deleted irrelevant information in the data, such as time, commenter’s name, etc.
- Removed special characters, expressions and stop words in both simplified and traditional Chinese characters
- Merged synonyms
- Calculated TF-IDF scores

2.3 Semantic network analysis

- We calculated betweenness centrality to determine the number of shortest paths passing through a point. The greater the number of shortest paths passing through a point, the higher its betweenness centrality [3].
- We calculated TF-IDF and betweenness centrality to observe words and nodes with relatively large influence. Clustering analysis shows us the clustered differences between comments from China and Taiwan on topics and helps to systematically understand the hierarchy of comments and composition of topics.

3. Results

3.1 Word cloud

The calculated TF-IDF scores are visualized in word clouds as shown in **Fig. 2** and **Fig. 3**.

Fig. 2. Word cloud based on Chinese TF-IDF scores

Fig. 3. Word cloud based on Taiwanese TF-IDF scores

3.2 Betweenness centrality analysis

Table 3. Betweenness Centrality (Chinese data)

NO.	Words	Betweenness centrality
1	global	73.495
2	China	60.747
3	America	49.619
4	progress	36.218
5	maximum	30.022
6	Japan	28.029
7	goods	24.122
8	Korea	23.588
9	good	23.025
10	significant	16.297

The Betweenness Centrality values of the Chinese data is shown in **Table 3**.

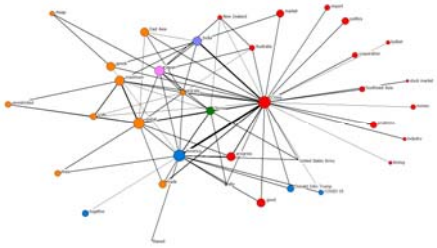


Fig. 4 Semantic Network diagram (Chinese data)

The network diagram in **Fig. 4** shows that red nodes centered around "China" take up half of the network. Orange nodes centered around "Global" represent the international influence, blue nodes "America", green nodes "Korea", pink nodes "Japan" and purple nodes represent comments that mention "India".

The calculated Betweenness Centrality values of Taiwanese comments are shown in **Table 4**.

Table 4. Betweenness Centrality of Taiwanese comments

NO.	Words	Betweenness centrality
1	Chinese mainland	28.197
2	America	24.345
3	world	24.301
4	people	24.199
5	Taiwan	19.978
6	problems	17.01
7	Japan	13.513
8	policy	13.491
9	life	13.354
10	COVID-19	13.238

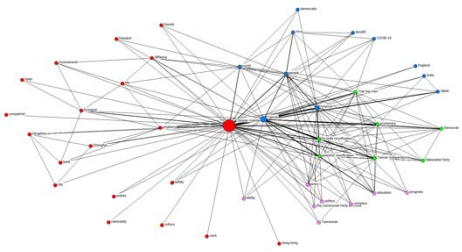


Fig. 5 Semantic Network diagram of Taiwanese data

As can be seen from **Fig. 5**, the red nodes centered around "China Mainland" take up half of the network. The blue nodes around "America" represent comments that mention the United States. The “Taiwan” node is located between "China Mainland" and "America".

3.3 Clustering analysis

While the previous analysis focused on nodes, the cluster analysis below can help us understand the composition of comments from the perspective of a "surface"[4].

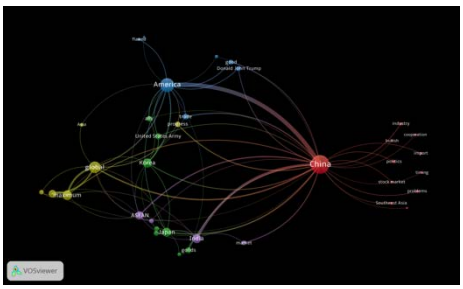


Fig. 6. Cluster diagram of Chinese comments

Table 5. Clusters of Chinese comments

NO.	Clustering	Word examples
1	Praise for China.	Bullish, China, cooperation, import, industry, money, politics, problems, Southeast Asia, stock market, timing
2	The cost of imports from America's Allies	Ally, Australia, cheap, goods, Japan, New Zealand, United States Army
3	The United States is still in an epidemic.	America, COVID-19, Donald John Trump, flawed, good, together, trade
4	RCEP's place in the world.	Asia, global, maximum, progress, scale, unrestricted
5	Discussions on ASEAN, East Asia and India markets.	ASEAN, East Aisa, India, market

- **Praise for China**

The first cluster in the Chinese data shows that Chinese netizens praise China for joining the pact. Keywords: *bullish, currency, import, industry, money, politics, the problems, the stock market, timing.*

- **Import tariffs from America's Allies decreased**

The second cluster groups comments that discuss the cost of import products that China imports from the United States. Keywords: *Ally, Australia, cheap, Goods, Japan, New Zealand, United States Army.*

- **The United States from the epidemic**

The third cluster groups comments about the United States being in the midst of an epidemic. Keywords: *America, COVID-19, Donald John Trump.*

- **RCEP and the world**

The fourth cluster groups comments about the influence of RCEP in the rest of the world. Keywords: *Asia, global, Maximum, progress, scale and unrestricted.*

- **Discussions on ASEAN, East Asia and India markets**

The fifth cluster groups comments that discuss ASEAN, East Asia and the Indian market. Keywords: *ASEAN, East Asia, India, market.*

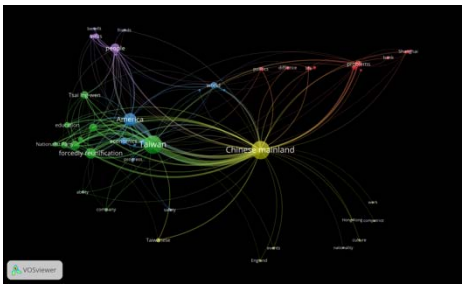


Fig. 7. Cluster diagram of Taiwanese data

Table 6. Clusters of Taiwanese comments

NO.	Clustering	Word examples
1	The Life experience of Taiwanese in China.	Bank, city, difference, foreigner, Hangzhou, hotel, inconvenient, life, politics, problems
2	Taiwan people's discussion of national reunification.	Ability, company, Democratic Progressive Party, education, forced reunification, Nationalist Party, peaceful reunification, policy, Taiwan, Taiwan independence, Tsai Ing-wen
3	Political and economic discussions with the United States, India and Japan.	America, democratic, economics, freedom, India, Japan, progress, safety, world
4	Cultural and nationality issues between China and Taiwan	Chinese mainland, compatriot, culture, England, events, nationality, Taiwanese, work
5	The Chinese Communist Party has benefited from the post-COVID-19 world landscape.	Benefit, COVID-19, friends, people, the Communist Party, virus

- **Life experience of Taiwanese in China**

The first cluster groups comments that include experiences of Taiwanese people living in a Chinese city. Keywords: *Bank, city, difference, Hangzhou, Hotel, Inconvenient, life, Politics, Problems.*

- **Taiwanese people's discussion of national reunification**

The second cluster groups comments on the issue of unification with China. Keywords: *Ability, Democratic Progressive Party, forced-party, Chen Chen, policy, Taiwan, Taiwan independence, Tsai ing-wen.*

- **Discussions regarding economy and politics in the United States, India and Japan**

The third cluster groups political and economic discussions regarding the United States, India, and Japan. Keywords: *America, democratic, economics, freedom, India, Japan, progress, safety, world.*

- **Cultural and identity issues between China and Taiwan**

The fourth cluster groups cultural and ethnic issues between China and Taiwan. Keywords: *Chinese mainland, compatriot, culture, events, nationality, Taiwanese.*

- **How the Chinese Communist Party benefits from COVID-19**

In the fifth cluster we find that Taiwanese netizens believe that the CCP gained benefits from the pandemic. Keywords: *Benefit, COVID-19, friends, people, the Communist Party, virus.*

3.4 Density analysis

Density maps help us to visually see the differences in number of comments by using brightness to resemble the amount of comments.

As shown in **Fig. 8**, most comments in the Chinese dataset mention "China", followed by "United States", "Global", "Japan", "ASEAN", "India", etc.

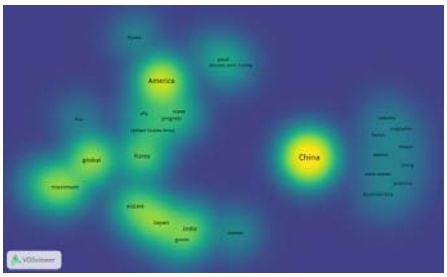


Fig. 8. Density diagram of Chinese comments

As shown in **Fig. 9**, most comments in the Taiwanese dataset mention "Mainland China", followed by "Taiwan", "The United States" and so on.

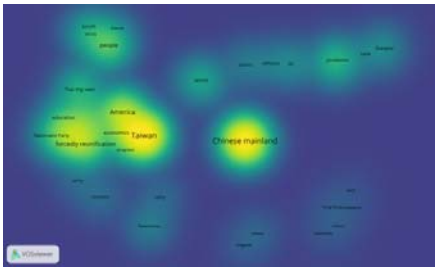


Fig. 9. Density diagram of Taiwanese comments

4. Conclusions

This paper applied semantic network analysis on comments crawled from Chinese and Taiwanese platforms that discuss the recent signing of the RCEP and showed differences in perception on the pact from netizens in both countries.

4.1 The United States and India have a strong presence

TF-IDF scores showed us the importance of topics among netizens in China and Taiwan that comment on the recent established RCEP. While the US and India are not included is this pact, both countries scored high in the list of topics showing that, despite being no participants, attention for both countries cannot be ignored.

4.2 Taiwanese netizens pay more attention to China and the United States than other topics

The betweenness-centrality scores of the dataset that included Taiwanese comments show that most topics discussed are focused on China and the United States.

4.3 Differences in topics between Chinese and Taiwanese netizens

Table 7. Comparison of topics in Chinese and Taiwanese comments

Market	1:Praise for China. 2:Discussions on ASEAN, East Asia and India markets.	
Import	The cost of imports from America's Allies fell.	
America	The United States is still in an epidemic.	Political and economic
Rcep	RCEP's place in the world.	
China		1：The Life experience of Taiwanese in China. 2：Taiwan people's discussion of national reunification. 3：Cultural and nationality issues between China and Taiwan 4：The Chinese Communist Party has benefited

- **The differences are significant**

From **Table 7** we can see that the topics discussed by Chinese and Taiwanese netizens vary greatly. It is remarkable that four out of the five most discussed topics in Taiwan mention China and the differences in life, unification, cultural and ethnic issues, and benefits for the Chinese Communist Party.

- **A common topic is the United States**

The only concern shared by the two groups of netizens are the United States of America. While Chinese netizens discussing the US are mostly concerned about COVID-19, the Taiwanese comments in this selection mostly focus on politics and the economy. In the perspective of netizens in both countries the influence of the US on the world economy seems to be fairly large.

- **Chinese netizens are more concerned about the economy**

Chinese comments focus on the influence on economic markets, such as the impact on China's domestic market and markets in ASEAN, East Asia, and India. They noticed that import tariffs have decreased and seem to value economic development.

- **The main concern of Taiwanese netizens is China**

While the search terms used to crawl the comments did not include the term ‘China’, most Taiwanese comments seem to discuss China and related issues, rather than RCEP itself. This seems to be the biggest difference in topics between the two datasets; Chinese comments discussing the economic consequences of the pact, while Taiwanese netizens seem to be more concerned with the political outcomes.

- **Reference value**

This research provides insights in the perception of netizens on the newly shaped RCEP from both included and excluded countries. It will be helpful to understand and deal with the concerns that may raise among citizens from countries that are excluded from economic agreements in the future.

5. Limitations

Due to time constraints, the textual analysis in this paper is limited. Future research can use

survey-based data [5] for diving deeper into this topic and provide more extensive insights.

***Funding Statement*:** This work was supported by a National Research Foundation of Korea (NRF) (http://nrf.re.kr/eng/index) grant funded by the Korean government (NRF-2020R1A2C1014957).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] Wilson, J. D, "Mega-regional trade deals in the Asia-Pacific: Choosing between the TPP and RCEP?," *Journal of Contemporary Asia*, *45*(2), 345-353, 2015.

[2] Sirisuriya, D. S, "A comparative study on web scraping," 2015.

[3] Brandes, U, "A faster algorithm for betweenness centrality," *Journal of mathematical sociology*, *25*(2), 163-177, 2001.

[4] Shi, J.-g., Miao, W., & Si, H, "Visualization and analysis of mapping knowledge domain of urban vitality research," *Sustainability*, *11*(4), 988, 2019.

[5] Nan, D., Kim, Y., Park, M. H., & Kim, J. H. "What Motivates Users to Keep Using Social Mobile Payments?," Sustainability, 12.17 6878, 2020.

Effect of Building Characteristics and Temporal Changes on Fire Alarms

Gwanmuk Lim, Seoltae Ko and Keon Chul Park
Data Innovation Team, Seoul Digital Foundation
Seoul, South Korea
[e-mail: realistic8997@sdf.seoul.k, kost@sdf.seoul.k, parkkc07@sdf.seoul.kr]
*Corresponding author: Keon Chul Park

Abstract

In the event of an emergency such as a fire, the installation of automated fire detection facilities that can promptly evacuate and respond to the residents is increasing. Can be lowered. Therefore, it is necessary to improve the system through efficient and periodic inspections and building environment surveys by finding the main factors that affect the occurrence of fire alarms. Based on this problem, this study attempted to find the influencing factors based on the current status of real-time fire alarms recorded by the IoT firefighting system of the Seoul Fire & Disaster Headquarters. As a result of the analysis, factors related to the scale, such as the total floor area of the building, had the greatest influence on the amplification of fire alarm generation. In addition, it was found that the occurrence of fire alarms increased in private buildings, buildings with R-type receivers, and buildings with a large number of failure days or alarm blocking days. In addition, as a result of looking at the factors influencing temporal and seasonal changes through the graph, it was found that they almost followed the patterns of daily activities of people during weekdays and weekdays, and the highest values were recorded around 10 am and 2 pm. Through the research, it was announced that it was necessary to investigate the building environment causing the occurrence of fire alarms, along with the system internal inspection, and to propose a method to increase the accuracy of the alarm by recording additional building environment data in real time.

***Keywords*:** Fire Detecting System, Fire Alarm, Environment of Building

1. Introduction

In the event of a fire, the rapid spread of emergency situations and initial response are very important measures to prevent human damage and minimize material loss. However, since there is a limit to humans' awareness of and propagation of an emergency situation, an automated fire detection facility is needed as a fire fighting system to take over the role. Recently, as the number of high-rise/complex buildings has increased and existing buildings have been advanced, the number of buildings equipped with automated fire detection facilities is increasing, and the importance thereof is also increasing. However, because the firefighting system detects the situation with a mechanical algorithm, there may be cases where the situation is mistaken for a “fire” even if it is not a fire. In other words, not only malfunctions due to failures such as system internal failures, but also non-fire alarm situations in which the system operates by erroneously detecting the situation even if it is not a fire may occur. As the number of buildings installing and building systems increases, the number of system malfunctions and the resulting accidental movements continue to increase proportionally. According to statistics on the current status of fire accidents in Seoul, the number of accidental exits in the last three

\

years has been clearly increasing.

Table 1 . Misunderstanding of fires in Seoul

	2016	2017	2018	2019
Number of Misleading Move	119	162	307	364
Increase Rate of Year (%)		0.36	0.90	0.19

Source: Seoul Open Data Plaza

Frequent occurrence of error signals from the automated fire detection facility may increase unnecessary manpower consumption and fatigue of the competent fire department, and may make it difficult to cope with the situation when an actual fire occurs at other points in the same time period. Resident workers in the building may experience inconvenience due to frequent fire alarms and do not trust the fire alarm. If the alarm does not sound due to power cut off of resident workers, etc. It can cause damage [2].

As the problem of the automatic fire detection facility is continuously raised, the internal defects of the system are continuously improved and the accuracy is improved through various technical studies and inspections. However, since the automatic fire detection facility may react sensitively depending on various external factors such as the internal environment of the building, its main purpose, and structural characteristics, it is considered necessary to study the system operation suitable for the building characteristics and the internal environment.

Therefore, research on improvement of internal error in the firefighting system should be conducted in parallel with research to understand how the automatic fire detection facility reacts to changes in the structural characteristics and environment of the building. Through this, it is thought that it will be worth using it as a basis for checking the main inspection items and efficient and intensive inspection time of the automatic fire detection facility.

2. Research Design

Based on the above background, this study analyzes the fire alarm record data of the real-time firefighting facility management system of the Seoul Metropolitan Fire & Disaster Headquarters to find out the correlation between the building characteristics and the number of fire alarms. In other words, the basic purpose of the study is to inform that frequent malfunctions of automated fire detection facilities may occur depending on the characteristics of buildings and firefighting systems and seasonal and time factors.

2.1 Data

The data used for the analysis is the fire alarm occurrence record data recorded in the real-time firefighting facility management system provided by the Seoul Metropolitan Fire & Disaster Headquarters. In this system, the contents of fire receiver events, including the occurrence of fire alarms in buildings designated as “specific fire targets,” are recorded in real time. In this analysis, a total of 714 fire-fighting target-designated buildings with meaningful records collected for 9 months from December 2019 to August 2020 were targeted.

2.2 Methodology

The data used for the analysis is the fire alarm occurrence record data recorded in the real-time firefighting facility management system provided by the Seoul Metropolitan Fire & Disaster Headquarters. In this system, the contents of fire receiver events, including the occurrence of fire alarms in buildings designated as “specific fire targets,” are recorded in real time. In this analysis, a total of 714 fire-fighting target-designated buildings with meaningful records collected for 9 months from December 2019 to August 2020 were targeted.

The order and specific contents of the study are as follows. First, basic statistics of major variables including the number of fire alarms are calculated to identify individual characteristics of each variable. Second, the characteristics of buildings and systems that have influenced the occurrence of fire alarms are confirmed by regression analysis. Public institutions, floor composition, number of residents, number of occupants, inspection days, breakdown days, alarm clock shut-off days, power alarm days, etc. are used as independent variables, and regression analysis is performed separately for the main purpose of the building. In addition, by grasping the current status of fire alarm occurrences by

month, day of the week, and time slot, we seek to find the optimal inspection time to help the firefighting system management.

2.3 Describe of the Variables

The main variables used in this analysis were largely divided into the characteristics of the building and the firefighting system (automatic fire detection facility). The total descriptive statistics of the main variable is that the number of fire alarm occurrences is 119.82 times and the maximum value is 1,718 times, compared to the average of 43.74 times and the median value of 9 times. Can be inferred. Because the deviation of other influencing factors is also quite large, it can be confirmed that both the building characteristics and the mechanical characteristics of the automatic re-detection equipment are individually different.

Table 2 Descriptive of the Variables

	mean	std	min	median	max
Number of Fire Alarms	43.74	119.82	0	9	1,718
Ground Floors	5.71	7.387	0	4	55
Basement Floors	2.20	1.682	0	2	9
Total Area of Building	14,790.16	21,245.9	0	9,122.25	257,422
Resident Number of people	402.32	780.44	1	137.5	10,102
Maximum Capacity Number of People	3,344.60	6,808.92	6	1,488	80,300
Number of Fire Detection Systems	321.70	885.386	1	122	15,683
Number of Inspection Days	2.44	5.936	0	1	63
Number of Breakdown Days	0.93	7.707	0	0	127
Number of Alarm Blocking Days	2.56	13.36	0	0	153
Number of Power Alarm Days	2.30	3.90	0	1	24

3. Result of the Analysis

3.1 Correlation of Area and Fire Alarms

The dependent variable that is the core of this study is the number of fire alarm occurrences. Since the number of fire alarm occurrences is recorded through fire detection by individual detectors, there is a possibility that the number of times increases significantly as the number of installations increases. Also, the fact that the number of detectors installed in the automatic fire detection facility is large means that the building is large. Therefore, the total number of fire alarms will inevitably be greatly influenced by the size of the building, such as the total floor area.

In this way, variables related to the building size (total area) and the number of automatic fire detection facilities that have a significant effect on the fire alarm are controlled in advance, or the records of alarms that have sounded simultaneously for a certain period of time are redefined and the variables are redefined, and other influencing factors will be examined. It was determined that it was necessary.

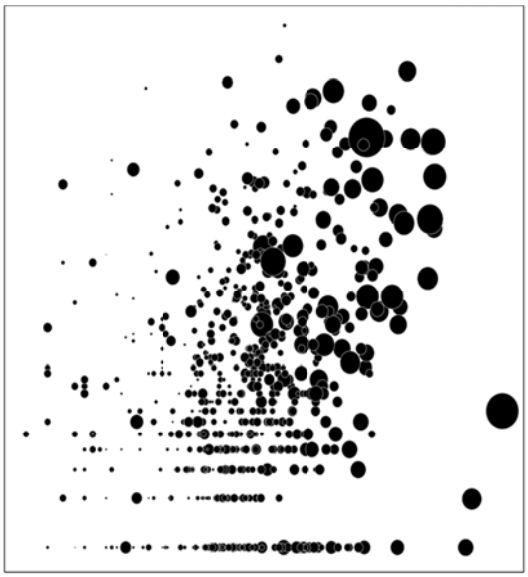


Fig. 1. Scatter plot of correlation between total floor area-number of automatic fire detection facilities-number of fire alarm occurrences (Size : Total Area / X : Number of Fire Detecting System / Y : Fire Alarms)

3.2 Effects of the Character of Buildings and Fire Fighting Systems on Fire Alarms

Next, the effect of the main characteristics of the building on the number of fire alarm occurrences was examined through multiple regression analysis. Multiple regression analysis is used to relatively compare the influence of each variable on the dependent variable when there are several independent variables and to determine the exact causal relationship between variables. [3]

Prior to the analysis, based on the analysis result that the size of the building has a great influence on the occurrence of fire alarms, the number of fire alarm occurrences is divided by the total building area to control the scale factor, and the number of fire alarm occurrences expressed by standardizing this variable is a dependent variable. Multiple regression analysis was performed by defining as. In addition, after performing the overall regression analysis, the data were divided according to the type of receiver and the use of the building to examine the detailed differences in the influencing factors.

As a result of analyzing the entire data, when the type of receiver installed in the building is R-type (within significance level of 0.05), the fewer basement floors, the more failure days and alarm block days, the fewer inspection days (within significance level 0.05)) It was found that the number of fire alarm occurrences increased. In addition, it was found that the number of fire alarm occurrences was higher for private buildings, not public institutions, which was significant at the 0.05 level.

Table 3. Result of Regression of Fire Alarms

(N=714)	β	S.E.	t
(Constant)	0.1234	0.120	1.030
Public Building	-0.3016*	0.124	-2.424
Detector_R Type	0.3378**	0.096	3.529
Ground Floors	0.0341	0.058	0.584
Basement Floors	-0.1127**	0.041	-2.721
Resident people	-0.0240	0.046	-0.521
Capacity People	-0.0053	0.042	-0.127
Inspection Days	-0.0740*	0.036	-2.057
Breakdown Days	0.2384**	0.035	6.857
Alarm Blocking Days	0.1593**	0.036	4.463
Power Alarm Days	0.0219	0.038	0.579
R-Square : 0.199 / F Statistic : 15.840 / P-value : 0.000			

Next, groups were classified according to the purpose of the building and each regression analysis was attempted. First, it is a transportation facility with the most used buildings. Due to the characteristics of Seoul's transportation, which is composed mainly of buses and subways, it was confirmed that the majority of the objects covered by transportation facilities were subway stations. As a result of the analysis, when the receiver type is R type, the more ground floors and the fewer failure days, the more the fire alarms tend to occur. (R-Square : 0.235 / F Statistic : 6.347)

Second, the analysis was conducted only on buildings intended for educational research facilities. Educational research facilities include elementary, middle and high schools, universities and research institutes, libraries, training centers, and academies. As a result of the analysis, when the receiver type is R type and when the number of ground floors is relatively small, the occurrence of fire alarm tends to increase. (R-Square : 0.130 / F Statistic : 2.433)

Third is the analysis result of buildings classified as business facilities. Regarding the receiver type, as in the other uses, there was a significant difference in influence, and it was found that the number of occurrences of the R-type was greater than that of the P. In addition, it was found that the more the number of alarm blocking days, the more the number of fire alarms occurred. In addition, it was found that the more residents, the greater the number of fire alarms. (R-Square : 0.290 / F Statistic : 3.225)

The results of the regression analysis of buildings classified as complex buildings showed that the F value did not meet the significance level within 0.05, so the model was not significant. Since it is a building with a mixture of various uses and characteristics, it is interpreted that it is affected irregularly without a certain causal relationship even though the number of fire alarms is very high. (R-Square : 0.072 / F Statistic : 0.575)

3.3 Effects of Temporal Factors

Next, we tried to see the effect of temporal factors and their changes on the occurrence of fire alarms. This relationship was determined by first calculating the frequency table for the number of fire alarm occurrences, and grasping the pattern through a visualization graph based

on this.

Table 4. Result of regression of fire alarms by building purpose

building purpose	Transp-ortation	Educate & Study	Business	Com-plex
(Constant)	-0.0686**	-0.9267	-0.1462	0.1239
Public Building	0.0063	0.0121	-0.1166	-0.6514
Detector_R Type	0.0111**	1.3990**	0.5390*	0.8010
Ground Floors	0.0125**	-1.6619**	-0.1651	0.1175
Basement Floors	-0.0015	-0.4330	-0.1142	-0.2559
Resident people	0.0017	-0.1412	0.3428*	0.0628
Capacity People	-0.0004	-0.3332	0.0147	0.0404
Inspection Days	0.0002	-0.1851	-0.0941	-0.4943
Breakdown Days	-0.0194*	0.0963	-0.0623	-0.0957
Alarm Blocking Days	0.0075	-0.6156	0.2302**	0.0839
Power Alarm Days	0.0000	-0.0173	-0.0822	0.6129

First, looking at the number of monthly fire alarms, the number of fire alarms increased significantly in January 2020, when the system was maintained to some extent after December 2019, and the number of fire alarms gradually increased until June. And July-August again markedly decreases, which can be thought that fire alarms sounded less due to seasonal factors in summer (e.g. humidity).

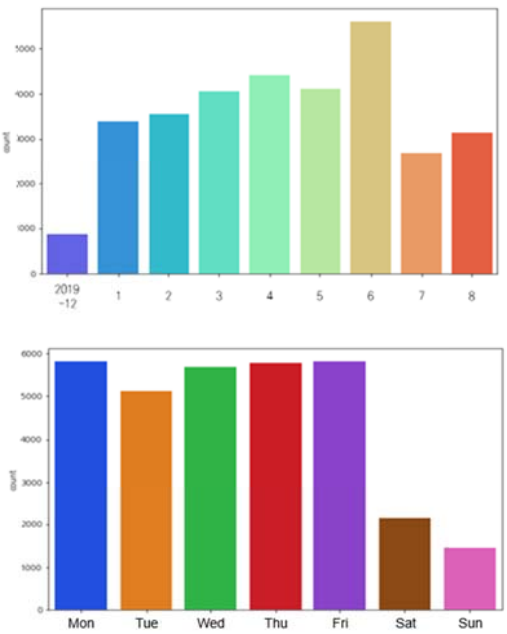


Fig. 2. Graph of number of fire alarm occurrences by month and day of the week

(Top: by month, bottom: by day of the week)

The result of visualizing the number of fire alarm occurrences by time is shown in Figure 5. The top left is the total fire alarm, and the following figure is a visualization of the graph by day of the week.

First of all, the number of occurrences is maintained at a low level from 0 o'clock to 4 a.m., and then gradually increases from 5 o'clock to 8 o'clock. After that, a steep increase was observed at 9 am, and the peak in the morning time was recorded around 10 am. Then, it decreases sharply from 11 o'clock, and the frequency appears very low for a moment at 12 o'clock in the daytime. And again, from 1 p.m., it shows a big increase, and it hits the highest in the afternoon around 2 p.m. After that, it showed a sharp decline after 3 pm, gradually decreased, and tended to appear with a relatively low number of occurrences after 6 pm.

All fire alarms were divided by day of the week and examined. For weekdays (Mon-Fri), there are differences in detailed values, but all of the graphs show similar patterns. On the other hand, in the case of weekends, it is relatively sporadic and irregular, and the total number of occurrences is less than on weekdays, except for the case of a temporary large increase at 10 am on Saturday and 3 pm on Sunday.

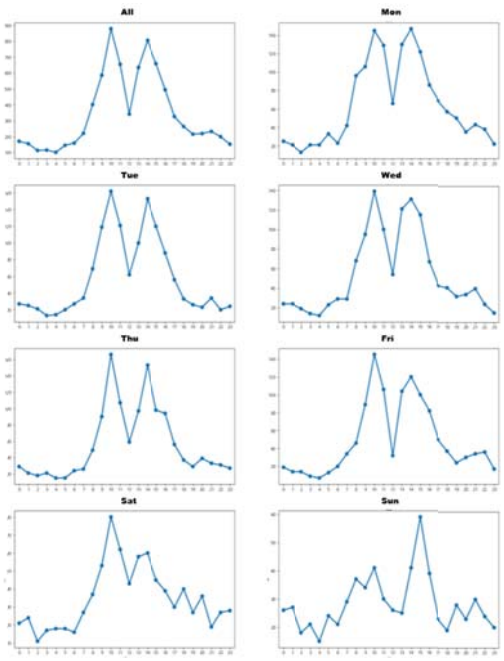


Fig. 3. Graph of number of fire alarm occurrences by time slot (total-day of the week)

The temporal pattern of fire alarms analyzed so far can be summarized as a result of following the flow of the time zone where people mainly engage in daily activities (economic activities, etc.) (weekdays 9:00 to 18:00: excluding lunch time). This results in similar results when the data is divided and analyzed according to the main use. Therefore, it can be seen that the automatic fire detection facility is sensitive to people's daily activities.

When the analysis was carried out by month, it was found that the overall number of occurrences appeared high in January-June when the system was stabilized, but decreased significantly in July-August. In addition to the internal maintenance factors of the system, it is believed that the seasonal factors of summer, including the long rainy season of July-August 2020, were at play.

In addition, the time when the highest frequency of alarm occurrence was shown was generally around 10~11 am and 2~3 pm on weekdays, which was a common result regardless of the use or seasonal factors. Therefore, it is expected that if the time for the inspection of the fire fighting system and the investigation of the environment inside the building is adjusted to this time, the factors influencing the occurrence of multiple occurrences of non-fire alarms can be diagnosed more efficiently and appropriate measures can be taken.

4. Conclusions

This study attempted to reveal that the occurrence and frequency of fire alarm malfunctions may be affected by the structural characteristics of the building, the mechanical characteristics of the fire detection facility (system characteristics), and other external influence factors. In order to solve the mechanical malfunction of the firefighting system and increase its reliability, it was emphasized through research that it is necessary to find the internal environment of the building or structural influencing factors, as well as the self-improvement of the firefighting system.

According to the results found through the study, it was found that the number of alarm occurrences was significantly increased even with a single fire because the amount of fire receivers installed was basically large depending on the size of the building such as the total floor area. In addition, it was confirmed that the occurrence and frequency of fire alarms may vary depending on the publicity and use of the building, the type of fire receiver and the frequency of failure and inspection, and the number of residents.

As further studies are conducted in the future and more data are collected, it is believed that a more detailed relationship between building characteristics and fire alarms can be identified. In addition, it is expected that the accumulated research results can be referenced even when additional automatic fire detection facilities are installed and operated.

References

[1] Hwang Hyeon-su, “Automatic Fire Detection Facility”, *Lighting and Electrical Equipment*, Vol. 9, No. 5, 1995.

[2] Jong-Hwa Lee, Chun-Ha Lee, Si-guk Kim, Ha-Sung Gong, “Measures to Reduce Non-Fire Insurance of Automatic Fire Detection Equipment”, *Journal of the Korean Society of Safety Management*, Vol. 13, No. 1, 2011.

[3] Doo-Sub Kim, Nam-Jun Kang, “Regression Analysis: Basics and Applications”, Nanam Publishing House, pp.85-116, 2014.

Before and after COVID-19: A news analysis of Tourism via semantic network analysis

Woomin Nam, Haesun Jung, and Jang Hyun Kim
Department of Applied Artificial Intelligence, Department of Human-Artificial Intelligence Interaction Sungkyunkwan University
Seoul, South Korea
[e-mail: applewoods@g.skku.edu (W.M.N.) , jestiriel@g.skku.edu (H.S.J.), alohakim@skku.edu (J.H.K.)]
*Corresponding author: Jang Hyun Kim

Abstract

The current study performed a semantic network analysis on the discourse regarding tourism industry using Naver news articles as the data set. In order to compare the discourse before and after COVID-19, the articles that included the keyword "여행(traveling)" in October 2019 and October 2020 were collected. For Data visualization and Semantic network analysis, NetworkX library was used. Results show that four clusters from 2019 and another four clusters from 2020 were different in that clusters generated for 2019 was mainly focused on travel, outside activities, and boycott of Japan whereas clusters created for 2020 were mainly focused on Corona and political issues. Contrary to previous expectations, the semantic network from the news articles were found to be more about social issues that are far from the content of the tourism industry's damage or individual inconveniences. The result seems to be due to the nature of the news media.

***Keywords*:** COVID-19, Corona Virus, Tour, Tourism, Travel industry, News, Pandemic, Semantic network

1. Introduction

Since the outbreak of the Corona virus in December 2019, it has become a serious problem all over the world. Despite different countries' effort to prevent the virus from spreading, the situation has built up to be a global pandemic. The pandemic has changed people's lives in many ways. Especially, it had a profound negative impact on travel-related businesses.

The spread of infectious disease can cause a major damage to the tourism industry. The increase in overseas travel, the expansion of cross-border exchanges, development of transportation can accelerate the spread. [1] Previous studies have shown that the spread of infectious diseases such as Severe Acute Respiratory Syndrome (SARS), avian influenza, and swine flu, have caused a sharp drop in demand from tourists, leading to a massive economic damage on the tourism industry. [2]

In case of airline industry, according to Kim (2020), the situation is very serious. It is estimated that the airline industry's sales forecast will drop by 68 percent year-on-year for several years, even after the COVID-19 crisis ends. [3]

According to Korea Economic Research Institute (KERI), the number of foreign tourists between March and June 2020 was 52,487. This

is a 99% drop compared to 5,105,686 from the same period 2019. It is speculated that this rapid decrease in the number of foreign tourists have led to the loss of 119,000 jobs in domestic tourism industry. It was also reported that the production inducement amount of the domestic tourism industry decreased by 13.2 trillion won. [4]

As of March 31, 2020, in case of Jeju Island, Korea's representative tourist destination, the number of tourists in Jeju Island dropped 58.4 percent year-on-year. Especially the number of foreign tourists decreased by 97.3 percent. The impact is expected to be greater compared to the 80.3 percent decrease in foreign tourists during the 8 months of Middle East Respiratory Syndrome (MERS) spread in 2015. [5]

As described earlier, Corona virus had a major negative impact on the tourism industry. Thus, the study seeks to investigate the impact of the Corona virus on the tourism industry through analyzing keywords in news articles from October 2019 and October 2020. The comparison can help us understand the impact of pandemics on human perception and life.

2. Literature review

1)Tourism industry and disease

The convenience of traveling has increased as travel has become more common due to the development of technology and globalization, but on the other hand, the spread of infectious diseases by travelers has begun to emerge. [6] In the case of SARS outbreak in 2003, China, Hong Kong, Singapore, and Vietnam suffered from more than $20 billion loss in GDP. The number of tourists also declined more than 70 percent across Asia. [7]

2) Pandemic

According to Jung (2020), 'Pandemic' refers to a situation when a new disease that people do not have immunity spreads around the world beyond expectations. Regardless of how serious the disease is, the degree of spread of the disease becomes a measure for the Pandemic declaration. The declaration of the pandemics means that the World Health Organization (WHO) and governments are not obliged to take any actions. However, the use of this term alone can trigger governments to invoke emergency procedures and arrangements, such as tighter travel ban and trade restrictions. The latest WHO case of pandemic was the swine flu outbreak in 2009. [8]

3. Methods

The present study collected Naver news articles that included the keyword '여행.' Collections were conducted on news articles from October 2019 and October 2020. The reason why these data were chosen was to compare the tourism before and after the Corona virus pandemic. 9,618 articles were collected for October 2020, and 7,411 articles were collected for October 2019.

After the collection, articles containing less than two travel-related keywords were removed, because they may be irrelevant to the actual tourism industry.

Table. 1. Data Description

	October 2019	October 2020
Total data collected	7,411	9,618
Data after cleaning	4,376	5,131

KoNLPy(Korean Natural language processing in Python) library was used for noun analysis. The data was preprocessed in the following order. First, words such as prepositions and conjunctions were deleted. Second, words with the same meaning were combined. Lastly, unnecessary words were deleted. After the refining process, top 1000 words sorted by Term Frequency – Inverse Document Frequency (TF-IDF) were used for the 2020 and 2019 data respectively. TF-IDF is the most used weight in information retrieval and text mining, a statistical figure that indicates how important a word is within a particular document.

Table. 2. Top 50 keywords (2019)

Word	TF - IDF	Word	TF - IDF
여행 (Travel)	179.45	서울 (Seoul)	69.77
일본 (Japan)	160.24	도시 (City)	69.06

관광 (Tourism)	156.07	증가 (Increase)	67.76
축제 (Festival)	149.31	고객 (Customer)	66.35
한국 (Korea)	110.51	투어 (Tour)	66.09
중국 (China)	108.15	여행객 (Traveler)	64.64
관광객 (Tourist)	103.82	국내 (Domestic)	63.82
지역 (Area)	100.82	참여 (Participation)	60.30
체험 (Experience)	96.01	기간 (Period)	57.95
문화 (Culture)	95.63	기업 (Corporation)	57.33
호텔 (Hotel)	92.87	한국인 (Korean)	57.23
가을 (Autumn)	87.62	콘텐츠 (Contents)	56.83
사진 (Photo)	87.59	이벤트 (Event)	56.57
서비스 (Service)	85.42	개최 (Hosting)	56.43
사업 (Business)	80.83	지원 (Support)	56.41
마을 (Village)	77.50	역사 (History)	55.19
상품 (Product)	77.01	홍보 (Advertisement)	54.22
공연 (Concert)	76.26	인천 (Incheon)	53.69
제주 (Jeju)	75.70	계획 (Plan)	52.70
프로그램 (Program)	75.22	관계자 (Officials)	51.52
노선 (Line)	74.40	예약 (Reservation)	50.86
세계 (World)	74.21	관광지 (Sight)	50.83
부산 (Busan)	74.00	해외 (Overseas)	50.50
방문 (Visit)	73.89	항공 (Flight)	49.25
사람 (Person)	71.36	판매 (Sales)	49.12

Table. 3. Top 50 keywords (2020)

Word	TF - IDF	Word	TF - IDF
여행 (Travel)	254.01	서울 (Seoul)	92.55
코로나 (Corona)	225.58	요트 (Yacht)	88.32
관광 (Tourism)	198.99	의원 (Congressman)	86.72
이일병교수 (Professor Lee Il-byeong)	175.00	세계 (World)	85.20
강경화 (Kang Kyung-wha)	150.05	계획 (Plan)	84.93
제주 (Jeju)	131.35	체험 (Experience)	84.50
국민 (Citizen)	127.86	공연 (Concert)	79.98
관광지 (Sight)	125.72	서비스 (Service)	79.92
사업 (Business)	125.42	외교부 (Minister of Foreign Affairs)	79.03
할인 (Discount)	125.19	가족 (Family)	77.79
방역 (Quarantine)	122.45	국내 (Domestic)	77.41
가을 (Autumn)	109.96	해외여행 (Trip abroad)	76.70
정부 (Government)	109.87	부산 (Busan)	74.35
지원 (Support)	107.46	프로그램 (Program)	72.83
사람 (Person)	106.74	소비 (Consumption)	72.82
상황 (Situation)	105.36	마을 (Village)	72.22
관광객 (Tourist)	104.18	해외 (Overseas)	70.15
지역 (Area)	103.64	상품 (Merchandise)	69.60
미국 (America)	101.66	출국 (Departure)	69.22
문화 (Culture)	99.82	영상 (Video)	69.17
한국 (Korea)	96.77	방문 (Visit)	68.95
중국 (China)	94.07	관광공사 (Tourist Service)	68.64
호텔 (Hotel)	93.07	재개 (Resume)	68.52
사진 (Photo)	92.94	공원 (Park)	67.51
온라인 (Online)	92.93	추석연휴 (Chuseok holiday)	66.66

Also, to identify the connectivity and degree centrality between keywords, Scikit-learn

Python library was used. For visualizing keywords, NetworkX library was used.

4. Results

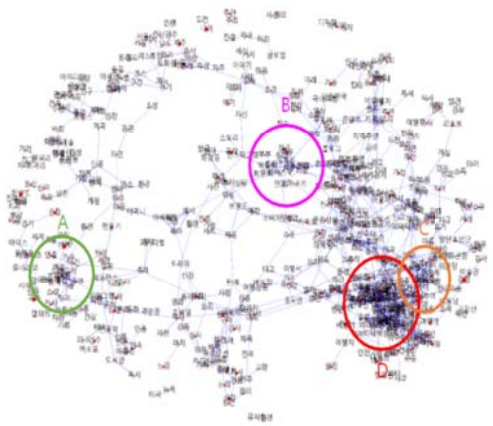


Fig. 1. Semantic Network of 2019 about tourism

Fig. 1 depicts a semantic network with keywords from 2019 articles. Here we extracted four meaningful clusters of A, B, C, and D. Cluster A is on cultural activities; "concert", "music", "art". Cluster B is composed of keywords related to job expansion in the tourism industry; "job", "recruitment", "tourism industry". Cluster C is composed of the keywords related to Japan boycott; "Japan government", "Boycott", "export regulation". Cluster D contains words related to travel around the world; "flight operation", "international Airport", "performance increase in airline industry".

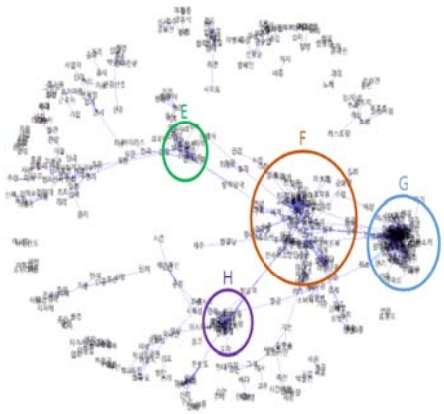


Fig. 2. Semantic Network of 2020 about tourism

Fig. 2 shows a semantic network made with keywords from 2020 news articles. In this case, overall graph could be divided into four groups: E, F, G, and H. Cluster E is made up of keywords related to COVID-19 group infections; "group infection", "confirmed case of corona virus", "gathering". Cluster F is the content of the Corona-related government policy; "quarantine rules", "Central Disaster and Safety Countermeasures Headquarters", "Social distancing". Cluster G is mainly composed of words about Professor Lee Il-byeong, the husband of Foreign Minister Kang Kyung-wha; "Professor Lee Il-byeong", "minister of foreign affairs and trade", "Yacht", "controversy". Cluster H are keywords related to sightseeing flight with the same destination and origin; "sightseeing flight", "Incheon International Airport", "flight".

Table. 4. Meaning of each cluster

Clusters from 2019 network		Clusters from 2020 network	
A	Keywords on Outdoor of cultural activities	E	COVID-19 group infections
B	Job expansion in the tourism industry	F	Government policy on Corona Virus
C	Boycott of Japanese in South Korea	G	About Political issue
D	Travel demand for countries around the world	H	Demand on the sightseeing tour

5. Conclusion

In the research design phase, it was expected that the semantic network from the news articles would show the changes in the tourism industry caused by COVID-19.

Contrary to previous expectations, the semantic network from the news articles were found to be more about social issues that are far from the content of the tourism industry's damage or individual inconveniences. The result seems to be due to the nature of the news media. The study sought to find notable keywords in news articles to find out the impact of the Corona virus on the tourism industry. The limitation of this study is the relatively short coverage of data collected. It only used data from October.

Studies based on this short period of time cannot represent the entire period. After the end of the COVID-19 crisis, further study based on the entire period from the beginning to the end of the COVID-19 outbreak can be suggested. As investigating public attitudes can be useful for specific industry to establish strategies [9], future research can also explore individuals' attitudes towards tourism during COVID-19 pandemic by employing other analysis such as sentiment analysis and computing the large-scale text data from social media. [10]

***Funding Statement*:** This work was supported by a National Research Foundation of Korea (NRF) (http://nrf.re.kr/eng/index) grant funded by the Korean government (NRF-2020R1A2C1014957).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] 홍민정, "코로나 19 확산에 대한 국내 잠재 관광객의 감정 반응 연구: 의미론적 네트워크 분석의 활용," *관광연구*, 35.3, pp. 47-65, 2020.

[2] Kuo, Hsiao-I., et al. "Assessing impacts of SARS and Avian Flu on international tourism demand to Asia," Tourism Management, 29.5, pp. 917-928, 2008.

[3] 김현일, "관광,항공산업 고사위기 직면...팬데믹 프리여권 등 도입해야." 동아일보 2020. Available:http://news.heraldcorp.com/view.php?ud=20201120000485

[4] KERI, "외국인관광객 급감으로 일자리 11.9만 개 증발 우려 <코로나19에 따른 외국인관광객 감소와 관광산업 유발 일자리 감소효과 추정>," Available:https://www.keri.org/web/www/news_02?p_p_id=EXT_BBS&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&_EXT_BBS_struts_action=%2Fext%2Fbbs%2Fview_message&_EXT_BBS_messageId=356044

[5] KIET, "코로나19가 국내 산업에 미치는 영향 Ⅱ. 코로나19가 서비스업에 미치는 영향과 대응방안," 2020. Available:http://www.kiet.re.kr/part/sDownload.jsp?s_idx=48098

[6] Findlater, Aidan, and Isaac I. Bogoch. "Human mobility and the global spread of infectious diseases: a focus on air travel," *Trends in parasitology*, 34.9, pp. 772-783, 2018.

[7] McKercher, Bob, and Kaye Chon. "The over-reaction to SARS and the collapse of Asian tourism," *Annals of tourism research*, 31.3 716. 2004.

[8] "코로나 19 세계적 대유행 '팬데믹'이란 무엇입니까?," Available:https://news.kbs.co.kr/news/view.do?ncd=4400004

[9] Nan, Dongyan, et al. "What Motivates Users to Keep Using Social Mobile Payments?," *Sustainability*, 12.17 6878, 2020.

[10] Jung, Jieun, et al. "When a girl awakened the world: A user and social message analysis of greta thunberg.," *Sustainability*, 12.7 ,2707, 2020.

KSII The 12th International Conference on Internet (ICONI) 2020.

Establishing a corpus for an AI-based empathic response system

Jae Hee Yang, Jin Yea Jang, Minyoung Jung, and Saim Shin
Artificial Intelligence Research Center, KETI
Sungnam-si, South Korea
[e-mail: {yangjaehee, jinyea.jang, minyoung.jung, sishin}@keti.re.kr]
*Corresponding author: Saim Shin

Abstract

The present study demonstrates the procedures for designing and developing a corpus in Korean, which consists of conversations between human users with depressive symptoms and an empathic response system. The conversations in the corpus were annotated according to seven basic emotions of the users (happiness, sadness, fear, surprise, anger, disgust, and neutral) and three responsive strategies of the system (clarification, empathic response, and suggestion). Specifically, the strategy of empathic response is carefully designed to include 10 distinctive response types (back-channel, facilitation, approval, disapproval, surprise, repeat, greeting, evaluation, encouragement, opinion) with five differing degrees of empathy. These annotations in the corpus may help train the system to detect various emotions of the users, to generate responses showing appropriate degrees of empathy, and to suggest activities that can help the users overcome their negative emotions when necessary.

***Keywords*:** Artificial intelligence, Corpus planning, Empathic response systems

1. Introduction

Recent years have seen growing attention on the importance of developing an AI-based empathic response system which can help people with depressive symptoms [1]. However, it is arguable that neither a successful system nor good quality datasets which can train such a system are available, especially for Korean users.

To provide viable datasets for training an empathic response system that can react to users with depressive symptoms, the present study takes the initial step of establishing a publicly available corpus in Korean, which contains conversations between users with depressive symptoms and an empathic response system. In what follows, the manner in which this corpus was designed and developed is illustrated.

2. The corpus

2.1 Planning

Under the consultation of two qualified psychiatrists in the field, typical features found in the dialogues of depressive people, key counseling strategies, and lists of mood-boosting activities were collected for reference Moreover, 842 therapeutic conversations of 50 minutes in length were analyzed, and Korean vocabulary items of emotions were compiled [2].

2.2 Design

In order for the response system to react to users with varying degrees of empathy, the present study took the following issues into consideration in the process of corpus design: (i)

This study was supported by Institute for Information & Communications Technology Promotion (IITP) and National IT Industry Promotion Agency (NIPA) grant funded by the Korea government (2017-0-00255 and S1601-20-1034).

the kinds of emotions the system is required to detect; (ii) the kinds of strategies the system uses to maintain the flow of conversations with users; and (iii) the forms of empathic responses the system uses to respond.

As for the types of emotions, in addition to Ekman's six basic emotions (happiness, sadness, fear, surprise, anger, and disgust), the category of 'neutral' was added for utterances which express a neutral state of emotions [3].

Concerning the system's responsive strategies, the present study employed three strategies widely-used in the clinical field (clarification, empathic response, and suggestion). **Table 1** shows examples of each strategy.

Table 1. Examples of the system's responsive strategies (Italics indicate Korean-English translations)

Strategy	Example
Clarification	*Are you in a sad mood now?*
Empathic response	*It must be difficult.*
Suggestion	*Why don't you go for a walk?*

Specifically, for the empathic response strategy, the present study suggests 10 distinctive response types with five differing degrees of empathy level by modifying and improving the empathic response types available from previous studies (Level 1: low, Level 5: high) [4,5]. **Table 2** illustrates these types with varying degrees of empathy, along with examples of each.

Table 2. Types and examples of empathic responses (Italics indicate Korean-English translations)

Response Type (Empathy Level)	Example
Back-channel (1)	*Yes./Okay.*
Facilitation (2)	*I want to hear more about it.*
Approval (3)	*That's a great idea.Go ahead.*
Disapproval (3)	*Please give it a second thought.*
Surprise (3)	*Really?*
Encouragement (4)	*Keep your chin up.*
Evaluation (4)	*That's harsh.*
Repeat (4)	*A: I'm going on a trip. B: A trip?*
Greeting (4)	*Good morning.*
Opinion (5)	*Do not skip your meal.*

2.3 Datasets

To user-initiated utterances, the system clarifies the users' emotional state or responds using one of 10 types of empathic responses. Additionally, suggestions for ways to recover from negative emotions are provided by the system when users agree to receive them.

Following this flow, two experts have been writing down conversations with annotations pointing out the user's emotion type, the system's response strategy, the system's empathic response type, and empathy level. These conversations will be released publicly when more than 25,000 items have been gathered in the datasets.

3. Conclusions

The present paper presents the elements taken into consideration when establishing a corpus for the development of an AI-based empathic response system. Further studies could be carried out to confirm or test the viability of the annotations suggested in this study. Using an emotionally supportive system trained with this corpus, ways to develop an online service which can mediate depression-related behaviors or have psychiatrists observe their patients' emotional changes in daily life can be sought.

References

[1] P. Kataria, K. Road, A, Jain, P. Dwivedi, and S. Bhingakar, "User adaptive chatbot for mitigating depression," *International Journal of Pure and Applied Mathematics*, Vol 118, No.16, pp.349-361, 2018.

[2] S.J. Sohn, M.S. Park, J.E. Park, and J.H., Sohn, "Korean emotion vocabulary and categorization of feeling words," S*cience of Emotion & Sensibility*, Vol 15, No.1, pp. 105-120, 2012.

[3] P. Ekman, "An argument for basic emotions," *Cognition & Emotion*, Vol 6, No. 3, pp. 169-200, 1992.

[4] K. Ito, M. Murata, T. Ohno, and S. Matsubara, "Relation between degree of empathy for narrative speech and type of responsive utterance in attentive listening," in *Proc. of the 12th Conference on Language Resources and Evaluation (LREC)*, 2020.

[5] H.W. Lee, "A study on empathic listener response strategies for promoting communication: Focusing on Korean language learners," *Teaching Korean as a Foreign Language*, Vol 42, pp. 241-269, 2015.

KSII The 12th International Conference on Internet (ICONI) 2020.

Strategic Dialogue Generation Model Through Multi-Task Learning

Jin Yea Jang, Minyoung Jung, and Saim Shin
Artificial Intelligence Research Center, KETI
Sungnam-si, South Korea
[{jinyea.jang, minyoung.jung, sishin}@keti.re.kr]
*Corresponding author : Saim Shin

Abstract

Due to the advancement of high-performance language models that are pre-trained with large amounts of text data, the performance of tasks in various natural language processing fields is also improving. Research on the development of a dialogue model that automatically generates a response sentence for a dialogue context using a pre-trained language model has also been actively conducted. In this paper, we introduce a dialogue generation model through multi-task learning based on a pre-trained language model. In the training phase, the dialogue model learns by performing both strategy prediction and response generation at the same time. The evaluation of the dialogue model utilizes the existing automatic evaluation metric and compares the model with the baseline model that simply trains the task for response generation. Experimental results show that the performance of the model through multi-task learning outperforms the baseline model and the use of dialogue history may contribute to the dialogue generation.

***Keywords*:** Artificial intelligence, Dialogue model, Language model, Multitask learning

1. Introduction

With the recent development of deep learning technology, dialogue generation models with various structures have been introduced. Multi-task learning is a technology in which a deep learning network simultaneously trains several tasks at the same time [1]. This was also used to create dialogues using personas [2]. In this paper, we implemented a dialog response generation model using multitask learning of two tasks (i.e., dialog response generation and dialog strategy prediction) and compared the performance with the baseline model. The paper proceeds in the order of introducing the experimental dataset, explaining the structure of the dialogue model, and presenting the experimental results.

2. Experiments

2.1 Dataset Development

We built the empathic dialogue dataset for the purpose of implementing service of emphatic dialogue. The dataset consists of user utterances and system responses according to three types of empathic dialogue strategies. **Table 1** shows examples of system responses. The size of the data is 4,905 dialogue pairs.

Table 1. Examples of the system's responsive strategies (Italic: translated into English from Korean).

Strategy	Example
Clarification	*Did you feel bad about it?*
Empathic response	*Oh my gosh!*
Suggestion	*Why don't you go for a walk?*

This study was supported by Institute for Information & Communications Technology Promotion (IITP) and National IT Industry Promotion Agency (NIPA) grant funded by the Korea government (2017-0-00255 and S1601-20-1034).

2.2 Multi-Task Model

We used a deep learning model that trains dialogue generation and strategy prediction at the same time to generate a response that matches the dialogue strategy for the user's utterance. **Fig. 1** illustrates the model architecture. The model has two heads in the Transformer Decoder structure: a language model head for dialogue generation and a classification head for predicting dialogue strategy. The model input is composed of either a combination of user utterance and system response or that of user utterance and distractor. Here, the distractor denotes a system response with a random dialogue strategy (not the true strategy of the system response) corresponding to the user's utterance. The language model head predicts the next token based on the generated tokens. The classification head predicts whether the system response of input is from a pair that matches the actual user's utterance or whether it is a distractor.

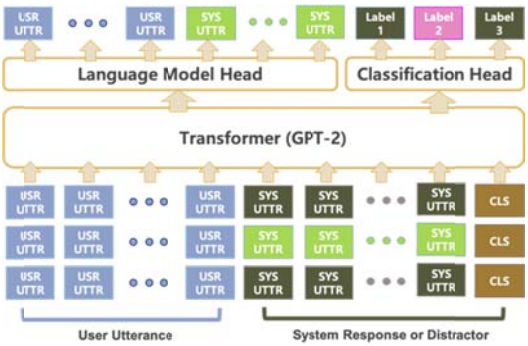


Fig. 1. Architecture of the multi-task dialogue model.

We performed fine-tuning based on the pre-trained language model, KoGPT-2 [3]. In the fine-tuning step, the total loss, which is the sum of the loss calculated from the two heads, is used for backpropagation. The data were divided into train, validation, and test sets at a ratio of 3:1:1. The optimizer used for training was AdamW [4], the epoch was set to 20, and the learning rate was 0.0001.

2.3 Results

We used a model with only a language model head as a baseline. In addition, we trained the model that added the dialogue history to the input in order to see the influence on the use of the dialogue history. A total of four different models were trained and the list is as follows.

- Single Task Learning (STL) is a model with a language model head.
- STL+H is a STL with the dialogue history information (H) added.
- Mult-Task Learning (MTL) is a model with both a language model head and a dialogue strategy classification head.
- MTL+H is a MTL with the dialogue history information (H) added.

Table. 2. Results of various evaluation metrics for system response

Metric	STL	STL+H	MTL	MTL+H
BLEU-1	16.54	13.27	**23.62**	22.08
BLEU-2	2.87	4.10	**8.10**	8.08
BLEU-3	0.72	1.39	3.21	**3.81**
BLEU-4	0.22	0.56	2.14	**2.60**
METEOR	16.73	22.50	**25.44**	25.10
Rouge-1	22.30	19.88	**30.70**	29.44
Rouge-2	3.76	6.07	10.17	**10.58**
Rouge-3	0.90	2.04	3.75	**4.69**
Rouge-4	0.27	0.83	2.47	**3.15**
Rouge-L	20.17	16.73	**28.73**	27.14

Table 2 shows the results of various evaluation metrics for system response (i.e., BLEU [5], METEOR [6], and ROUGE [7]). The multi-task learning model showed better performances in all evaluation metrics than the single-task learning models. In the case of the models in which the dialogue history was added to the input, the result of the higher n-gram evaluation metrics showed generally better results than the models without the dialogue history.

3. Conclusions

In this paper, we implemented a dialogue generation model using multi-task learning that trains response generation and dialogue strategies for user utterance at the same time. Through experiments, we confirmed that our model showed better performance than the single task learning models. The results also showed that the use of dialogue history can help generate more accurate responses. In future work, we plan to compare the dialogue model linked with a separate dialogue strategy classifier and a multitask learning model. There are also plans for expanding the type of dialogue strategy.

References

[1] Caruana, Rich. "Multitask learning." Machine learning 28.1, 41-75, 1997.

[2] Wolf, Thomas, et al. "Transfertransfo: A transfer learning approach for neural network based conversational agents." arXiv preprint arXiv:1901.08149, 2019.

[3] https://github.com/SKT-AI/KoGPT2

[4] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980, 2014.

[5] Papineni, Kishore, et al. "BLEU: a method for automatic evaluation of machine translation." Proceedings of the 40th annual meeting of the Association for Computational Linguistics, 2002.

[6] Banerjee, Satanjeev, and Alon Lavie. "METEOR: An automatic metric for MT evaluation with improved correlation with human judgments." Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization, 2005.

[7] Lin, Chin-Yew. "Rouge: A package for automatic evaluation of summaries." Text summarization branches out, 2004.

An Efficient Software Design of Ornamental Fish Data Management for Quarantine Purposed Electronic Illustrated Book

Junsu Kim, Changho Han, Sungyoon Cho and Kiwon Kwon
Korea Electronics Technology Institute (KETI)
Seongnam-si, South Korea
[e-mail: {jesus0084, chris.han, sycho, kwonkw}@keti.re.kr]
*Corresponding author: Kiwon Kwon

Abstract

This paper proposes a methodology of application design that provides data collection and management for organization of various ornamental fish data. In a general sense, target data collection and its management are not clearly pre-defined during an application development process, and the corresponding data structure is frequently modified. Therefore, it is required to design a software that allows easy access to dynamic data structure and changes on the detailed items without significant modifications. In this paper, the user-friendly software for data collection and management of unfixed data structure is proposed for convenient user-level access that to be beneficial for reducing useless development efforts.

***Keywords*:** Data collection, data management, data structure methodology, user convenience

1. Introduction

As the industry of ornamental fish has grown worldwide[1], the number of ornamental fishes imported into Korea have been dramatically increased[2]. The imported fishes mandatorily experiences quarantine, where the corresponding process is being still conducted by hard copy references that make quarantine procedures tardy. To overcome this inefficient process, developing electronic illustrated ornamental fish book receives big attention for the process improvement. However, the data structure for the quarantine is not clearly pre-defined yet, and a certain fish has many different common names in different regions or countries. Eventually, these problems cause frequent modification of the data structure in the development process.

In this paper, an efficient software of ornamental fish data management for quarantine purposed electronic illustrated book is proposed and designed. The designed software provides to design the ornamental fish data structure dynamically and manage the fish data based on the dynamically designed structure.

2. Design

The electronic illustrated ornamental fish book is proposed for real-time quarantine process, which allows that quarantine officers would be able to access the software wirelessly and

This research was a part of the project titled 'Development of electronic illustrated book for ornamental fish', funded by the Ministry of Oceans and Fisheries, Korea.

simultaneously, and the data would be updated immediately. The proposed electronic illustrated book is designed by Ubuntu based web server system shown in **Fig. 1**. The ornamental fish data management software is implemented by Flutter which provides easy implementation on various platforms. The proposed software is operated under the structure of mapping manager and data manager, respectively.

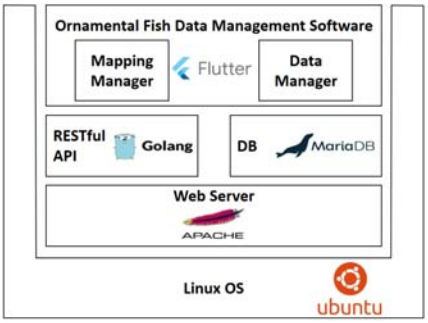


Fig. 1. Structure of Electronic Illustrated Book for Ornamental Fish

2.1 Mapping Manager

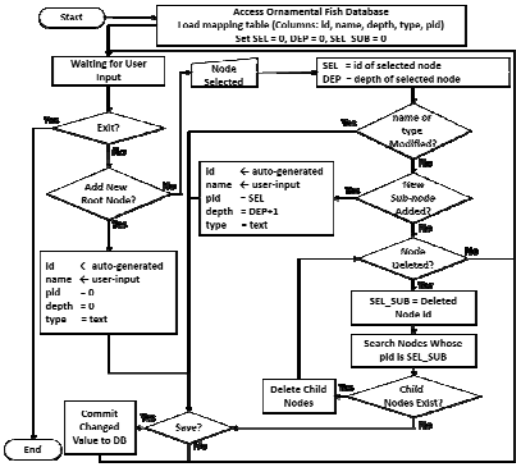


Fig. 2. Flowchart of Mapping Manager

The mapping manager illustrated in **Fig. 2** provides a method for users to design the ornamental fish data structure dynamically. In general, nodes of data structure are located on columns in database table, however, the mapping manger saves the nodes as data on the table called mapping table with column information such as id, name, depth, type, and pid. The column information indicates the identification, name, depth on tree, data type, and parent id of a node respectively.

2.2 Data Manager

The data manager illustrated in **Fig. 3** allows users to manage ornamental fish data based on the designed structure. Two tables are designed for managing ornamental fish data, fishdata table and fishname table. Each fish's data themselves are located on the fishdata table. The scientific name of ornamental fish is located on the fishname table so that the ornamental fish data can be classified according to its node and ornamental fish name. Therefore, the fishdata table has column information with node's id in the mapping table and scientific name's id in the fishname table as mappingid and fishnameid.

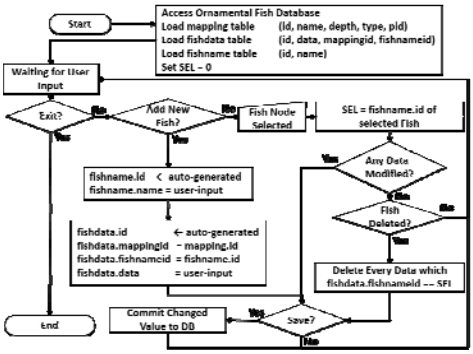


Fig. 3. Flowchart of Data Manager

3. Conclusions

In this paper, an efficient software of ornamental fish data management for quarantine purposed electronic illustrated book is proposed and designed. As the proposed software supports easy access and construction for flexible structure, the software complexity inevitably increased. However, it is expected that the designed software will provide efficient development process by reducing useless software revision efforts.

References

[1] R Loh, J Vukcevic and G Vastos Gomes, "Current status of aquatic veterinary services for ornamental fish in Australasia", *N Z Vet J*, Vol. 68, Issue 3, pp. 145-149, 2020.

[2] Dae-Young Kim and Jong-Ho Kang, "Improvement of Ornamental Fish Industry through analysis of recognition and market scale of the Ornamental Fish", *J Fish Bus Adm*, Vol. 43, No.3, pp. 89-106, 2012.

Speech Emotion Recognition on a Korean Multimodal Emotion Dataset Using Random Forest Classifier

Boseon Hong, Jin Yea Jang, Saim Shin, Hyedong Jung and Minyoung Jung*
Artificial Intelligence Research Center, Korea Electronics Technology Institute
Seongnam, South Korea
[e-mail: {bshong, jinyea.jang, sishin, hudson, minyoung.jung}@keti.re.kr]
*Corresponding author: Minyoung Jung

Abstract

Emotion recognition is an important task in various applications, and significant research effort has been spent on speech emotion recognition (SER). Despite the effort, SER is still challenging because emotions are subjective. Recently the Korean multimodal emotion dataset containing 10,339 audio-visual data has been published. Each data in the dataset is labeled by its corresponding emotion from seven emotions. We perform experiments of SER on the Korean dataset using random forest classifier.

***Keywords*:** Korean multimodal emotion dataset, speech emotion recognition, speech emotion classifier, machine learning, random forest

1. Introduction

Emotion recognition plays an important role in various applications such as robotics and human-computer interaction. Although much research effort has been spent on recognizing emotions from speech, speech emotion recognition (SER) is still a challenging task because emotions are subjective.

The well-known public datasets for SER are the English IEMOCAP [1] and the German Emo-DB [2]. The target emotions of the IEMOCAP are happiness, anger, sadness, frustration and neutral, whereas those of Emo-DB are happiness, anger, sadness, boredom, disgust, anxiety/fear and neutral. The Korean multimodal emotion dataset [3] has been recently introduced. The Korean dataset consists of audio-visual data, including facial recordings, speech and text transcriptions, and contains seven target emotions (happiness, anger, sadness, disgust, fear, surprise and neutral).

To the best of our knowledge, no research work for SER has utilized the Korean multimodal emotion dataset. We conduct experiments of SER on the Korean dataset using the traditional machine learning (ML) classifier, random forest (RF).

2. Dataset

The total 10,339 audio-visual data in the Korean multimodal emotion dataset were recorded from 100 Koreans who are acting aspirants or experts. They played according to the given transcription, emotion and circumstance description. **Table 1** shows the well-balanced data distribution by

This work was partly supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (2017-0-00255, Autonomous Intelligent Digital Companion Framework and Application) and National IT Industry Promotion Agency (NIPA) grant funded by the Korea government(MSIT) (S1601-20-1034, Development of artificial intelligence based empathy service technology for emotional stability).

emotions. We split the dataset into training and test sets in a ratio of 8:2.

Table 1. Data distribution by emotions

Emotion	Number of data
Happiness	1,492
Surprise	1,491
Neutral	1,477
Fear	1,471
Disgust	1,469
Anger	1,467
Sadness	1,472

3. Speech Emotion Classifier

3.1 Audio Feature Extraction

To extract audio features, we convert the audio-visual data (m2ts files) into the audio data (wav files). And we convert the audio signals into the digital signals at a 44.1kHz sampling rate. Then we extract eight audio features (the mean and standard deviation of the sampled data, the mean and standard deviation by calculating the root-mean-square for each frame, the number of counted parts when silent, the maximum and standard deviation of autocorrelation on center-clipped frames, and the high-pitch), and those features are input of the speech emotion classifier.

3.2 Random Forest Classifier

We adopt the RF classifier described in [4] for SER, because RF outperforms other traditional ML-based classifiers such as gradient boosting and support vector machines on the IEMOCAP dataset. RF creates several decision trees to solve the classification task. The number of decision trees used in our experiments is set to 1000.

4. Experimental Results

Table 2 displays performance of SER on the Korean dataset using RF classifier in terms of accuracy, f1-score, precision and recall. Compared to performance on the IEMOCAP dataset depicted in [4], one of reasons causes the performance gap between two datasets would be a different number of target emotions.

Fig. 1 illustrates the results of SER using RF classifier as a confusion matrix. We observed the data expressing sadness, anger, neutral and surprise are more accurately predicted than the data labeled as happiness, disgust and fear.

Table 2. Performance of RF classifier

	RF classifier
Accuracy	34.7%
F1-score	33.0%
Precision	32.4%
Recall	34.8%

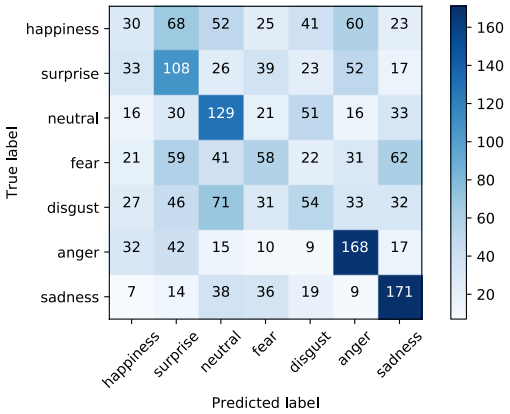


Fig. 1. Confusion matrix of RF classifier

5. Conclusions

Emotion recognition is a challenging and crucial task in many applications. In this paper, we conducted SER on the recently introduced Korean multimodal emotion dataset using RF classifier. The RF classifier showed better performance for sadness, anger, neutral and surprise emotions than happiness, disgust and fear emotions.

References

[1] Busso, C., Bulut, M., Lee, C. et al., "IEMOCAP: interactive emotional dyadic motion capture database," *Lang Resources and Evaluation*, Vol 42, No. 4, pp. 335-359, 2008.

[2] F. Burkhardt, A. Paeschke, M. Rolfes, W. F. Sendlmeier, and B. Weiss, "A database of german emotional speech," in *Proc. of Interspeech*, 2005.

[3] Korean multimodal emotion Dataset, "https://www.aihub.or.kr/keti_data_board/visual_intelligence," retrieved Sep. 2020.

[4] G. Sahu, "Multimodal speech emotion recognition and ambiguity resolution," *CoRR*, abs/1904.06022, 2019.

Identification of Trace TNT Peak Overlapped by Lactic Acid One Using Deconvolution Method in IMS Analysis

Sung-Seen Choi, Chae Eun Son, and Eunji Chae
Department of Chemistry, Sejong University,
Seoul, Republic of Korea
[e-mail: sschoi@sejong.ac.kr, ChaeunMCL@sju.ac.kr, codmswl9512@sju.ac.kr]
*Corresponding author: Sung-Seen Choi

Abstract

This paper presents a peak deconvolution method for identification of trace 2,4,6-trinitrotoluene (TNT) in the presence of lactic acid in ion mobility spectrometric analysis. In IMS spectrum, peak of TNT was overlapped to some degree with that of lactic acid, and it was covered by lactic acid peak when amount of TNT is much less than that of lactic acid. The overlapped peak was smoothed and deconvoluted to identify the presence of TNT. The ion drift times (**t_d**s) of the deconvoluted peaks of lactic acid and TNT were slightly changed. The **t_d** of lactic acid became longer whereas that of TNT became shorter.

***Keywords*:** Ion mobility spectrometry, 2,4,6-trinitrotoluene, lactic acid, peak devonvolution

1. Introduction

By increasing terror threat, security programs have been reinforced and detection of explosives is very impoprtant for security applications. Ion mobility spectrometer (IMS) is one of portable and hand-held explosive detectors, and it is used by security personnel to detect the presence of explosives in vehicles, packages, and other items [1-6]. Most organic explosives contain nitro groups and can be divided into nitroaromatic compounds, nitrate esters, and nitroamines. 2,4,6-Trinitrotoluene (TNT) belong to the nitroaromatic compounds and it is one of the most popular explosives.

For IMS analysis, trace explosive materials is generally collected with a smear matrix by rubbing hands or on the surface of objects. If the objects are touched with hands, the surfaces are contaminated with sweat. When rubbing the surface, some sweat must be collected along with trace explosive materials. Lactic acid is one of major sweat components, and its peak in IMS spectrum is overlapped to some degree with TNT one because difference in their ion drift times (**t_d**s) is not big and the FWHMs are relatively big. In the present work, we tried to identify the trace TNT peak from the overlapped peak of lactic acid and TNT in IMS spectrum using peak deconvolution method.

2. Experimental

2.1 Sample Preparation

TNT and lactic acid solutions were prepared by dissolving each explosive in acetone. Concentrations of the sample solutions were controlled by diluting the initial sample solution of 100 μg/mL. PTFE sheet (supplied by Newone S&T, Republic of Korea) was used as a smear matrix. Sample solution was dropped on the smear matrix and then the solvent was evaporated at room temperature before

introducing the smear matrix into the sample inlet part of IMS. The mixture sample of lactic acid and TNT was prepared by dropping the lactic acid solution and next dropping the TNT one on the smear matrix.

2.2 IMS Analysis

Portable IMS equipped with corona discharge ionization source, Ionab IMS of Newone S&T (Republic of Korea) was used. The analysis conditions of IMS were as follows: temperature of the sample inlet region was 150°C, temperature of the drift tube was 100°C, the electric field was 200 V/cm, and the drift distance of ion was 12.3 cm. The IMS analysis was performed in the negative mode.

2.3 Peak Deconvolution

Raw IMS peak was smoothed, and the smoothed curve was deconvoluted to two peaks of lactic acid and TNT.

3. Results and Discussion

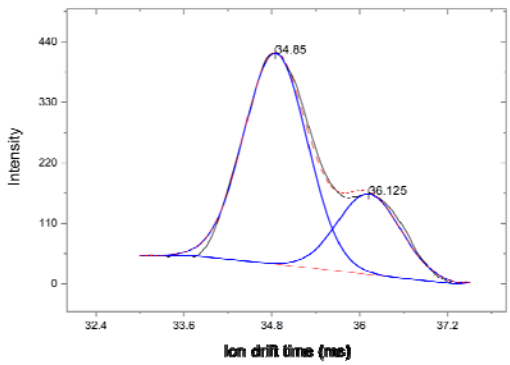


Fig. 1. Deconvolution result for IMS spectrum of mixture of TNT and lactic acid (1/10)

The $\mathbf{t_d}$s of lactic acid and TNT are 34.95 and 36.83 ms, respectively, when they are analyzed each other. Figure 1 shows IMS spectrum of the mixture of lactic acid and TNT with the peak height ratio of about 10/1 (lactic acid/TNT). Since two peaks were overlapped, it is hard to identify the TNT peak. As a result of the peak deconvolution, the overlapped peak can be separated to two peaks at 34.85 and 36.13 ms corresponding to lactic acid and TNT, respectively. Compared to the single sample $\mathbf{t_d}$s, the $\mathbf{t_d}$s of the deconvoluted peaks were slightly changed. The $\mathbf{t_d}$ of lactic acid became longer whereas that of TNT became shorter.

4. Conclusions

In this paper, we present identification method of trace TNT in the presence of lactic acid using peak deconvolution.

Acknowledgements

This work was conducted as a part of the research projects of "Development of automatic screening and hybrid detection system for hazardous material detecting in port container" financially (20200611) supported by the Ministry of Oceans and Fisheries, Republic of Korea.

References

[1] S.-S. Choi and C. E. Son, "Analytical method for estimation of transfer and detection efficiencies of solid state explosive using ion mobility spectrometry and smear matrix", *Analytical Methods*, Vol 9, No. 17, pp. 2505-2510, 2017.

[2] J. Lee, S. Park, S. G. Cho, E. M. Goh, S. Lee, S.-S. Koh, and J. Kim, "Analysis of explosives using corona discharge ionization combined with ion mobility spectrometry–mass spectrometry", *Talanta*, Vol 120, No. 1, pp. 64-70, 2014.

[3] A. Zalewska, W. Pawłowski, and W. Tomaszewski, "Limits of detection of explosives as determined with IMS and field asymmetric IMS vapour detectors", *Forensic Science International*, Vol. 226, No. 1, pp. 168-172, 2013.

[4] M. Najarro, M. E. D. Morris, M. E. Staymates, R. Fletcher, and G. Gillen, "Optimized thermal desorption for improved sensitivity in trace explosives detection by ion mobility spectrometry", *Analyst*, Vol 137, No. 11, pp. 2614-2622, 2012.

[5] K. M. Roscioli, E. Davis, W. F. Siems, A. Mariano, W. Su, S. K. Guharay, H. H. Hill, Jr., "Modular ion mobility spectrometer for explosives detection using corona ionization", *Analytical Chemistry*, Vol 83, No. 15, pp. 5965-5971, 2011.

[6] S.-S. Choi, O.-B. Kim, Y.-K. Kim, S. G. An, M.-W. Shin, S.-J. Maeng, and G. S. Choi, "Negative ion formation of pentaerythritol tetranitrate in atmospheric pressure chemical ionization-mass spectrometry and in corona discharge ionization-ion mobility spectrometry", *Bulletin of Korean Chemical Society,* Vol 32, No. 3, 1055-1058, 2011.

A Study on Periodically Log Broker in Cloud FaaS Applications

Dongmin Kim, Hyunwoo Kim, Kiman Jeon and Jae Gi Son
Department of Human IT Convergence Research Center, KETI
Gyeonggi-do, South Korea
[e-mail: dmkim@keti.re.kr, hwkim@keti.re.kr, kmjeon@keti.re.kr, jgson@keti.re.kr]
*Corresponding author: Dongmin Kim

Abstract

FaaS is a cloud-based implementation of microservices theory, a paradigm for a new software architecture in the FaaS(Function as a Service) style. A general open source based FaaS framework execute in the Kubernetes environment, and FaaS services depend on this environment. These are tools such as Prometheus and Heapster that support the monitoring function at the system operation level, but there is no structed tool for monitoring at the FaaS service application level. Therefore, in this study, we propose a log broker module structure that can observe logs generated at the FaaS-based application level. Through this study, the proposed log broker module has verified its function through interworking with Prometheus, and confirmed its applicability in a FaaS and edge application model that is sensitive to service operation time in the future.

***Keywords*:** Cloud monitoring, Edge log service, FaaS, Serverless Application

1. Introduction

FaaS is a cloud-based implementation of microservices theory, a paradigm for a new software architecture in the FaaS(Function as a service). FaaS is easy to use and saves resources. For this reason, FaaS is superior to legacy cloud computing and has a high interest in the cloud computing market. Existing open source FaaS platforms use the Kubernetes. Kubernetes-based FaaS platforms perform monitoring using a variety of open source tools for smooth operation.

Prometheus is a representative open source tool for monitoring metric data based on a pulling model. Prometheus provides a library for easy monitoring even when collecting Metric when does not support data types. Kubernetes' monitoring tools include cAdvisor and Heapster. These are monitoring tools that collect Metric such as CPU status, memory, network, and available disk space. Although they provide efficient monitoring on an open source basis, they are not suitable as tools for log collection at the actual FaaS application level because they only collect or manage log information at the system level.

There are also studies on log collection and analysis of docker and container-based applications. Kandan's research [1] proposed a framework for managing logs generated by container-based applications. In this study, we provided an orchestration framework that manages logs using a method for collecting or managing different types of log data in an environment where many container-based applications are deployed. Also, it introduced a

This work was supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No.2020-0-00985, Development of Lightweight System Software for Edge Server System Resource Management and Control)

tool that can be integrated with any container solution. SATNAM SINGH's work [2] established a logging infrastructure at the cluster level. This research introduces an easy logging system by abstracting the applications running in each container using existing open source tools. While these studies collect logs from FaaS applications, they provide extended capabilities over existing open source tools, but they do not provide log collection such as triggering time information for time series log information collection.

In this paper, we developed a log broker module for cloud-based FaaS applications. The log broker added a timer-based time trigger function to efficiently collect time series log information. The developed log broker verified the function through interworking with Prometheus and confirmed the applicability in the FaaS application model that supports the real time in the future.

2. Related Works

2.1 Kubernetes based FaaS architecture

Kubernetes is an automated tool for distributed management and operation of Docker containers developed by Google. Kubernetes provides scalability and optimal performance for container-based applications. Therefore, Kubernetes is well suited as an automation tool for cloud-based FaaS applications built on Docker containers. **Fig. 1** shows the architecture of a FaaS architecture based on Kubernetes.

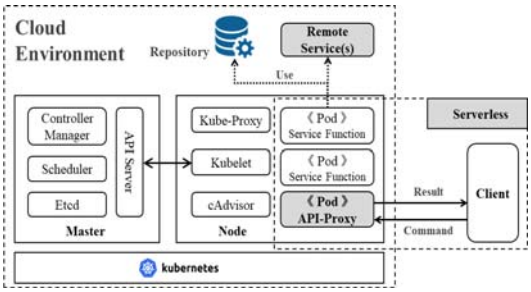


Fig 1. Architecture of Kubernetes-based FaaS

Kubernetes consists of a master and a node. The main components of Kubernetes for understanding this study are as follows.

- Pods: Pods are containers created in Kubernetes. It is the smallest unit of applications distributed to Kubernetes.

- Scheduler: Manages resources for pods running in Kubernetes and distributes pods to nodes.

- cAdvisor: Monitors containers running on pods. Only system level monitoring is possible.

Kubernetes-based FaaS applications are configured on Pods. In pods, API-Proxy communicates with clients who want to invoke FaaS applications. Also, it receives a command from the client and mediates a service through communication between internal pods.

2.2 Prometheus

Prometheus is an open source based tool that monitors metrics collected from where metric generated. The types of metric data include CPU and memory usage and http request time. Prometheus provides a suitable data format for collectable Metric. If there has no data format, Prometheus collects Metric using an exporter. Prometheus consists of a central server and an export node that generates Metric. This is a pulling model structure in which the central server collects Metric from export nodes. **Fig. 2** shows the structure of Prometheus.

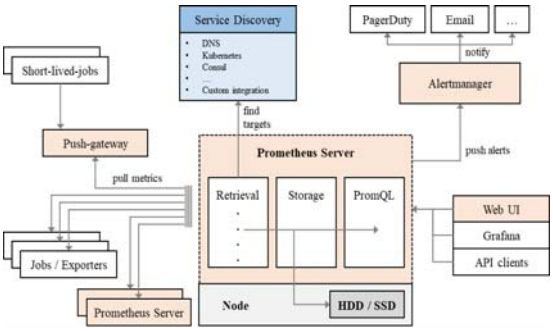


Fig 2. Structure of the Promethus

The main components of Prometheus are:

- Prometheus Server: The Prometheus server is the main element of Prometheus that collects metrics from export nodes.

- Web UI: The Web UI visualizes metrics in the form of text / html.

- Altermanager: In the event of a problem when collecting metrics, it acts as an alarm.

- push-gateway: Push-gateway is a kind of exporter that delivers metrics from export nodes directly to Prometheus servers.

3. Periodic Log Broker in FaaS Application

The log broker developed in this study delivers log or metrics to Prometheus' push-gateway periodically when log and metrics occurs in FaaS applications running in the cloud. Push-gateway can be used in this study as a way to compensate for the shortcomings of the pulling model.

First, a periodic execution model is required to generate periodic logs in the log broker. In this study, we design and utilize Periodic function manager (PF-Manager) that manages service functions based on timer in FaaS application running in FaaS. The designed PF-Manager manages functions through registration and execution phase. The registration phase registers the function with the name and time information of the function to be run periodically. The execution phase ensures the periodic operation of the functions of the FaaS application by calling the registered function periodically using a timer.

Fig. 3 shows the application structure using the PF-Manager and log broker in FaaS application. The log broker configures a function for generating a log in a FaaS application, and registers and executes the configured function by using the PF-Manager. The log generation function is to send log information to Prometheus' push-gateway.

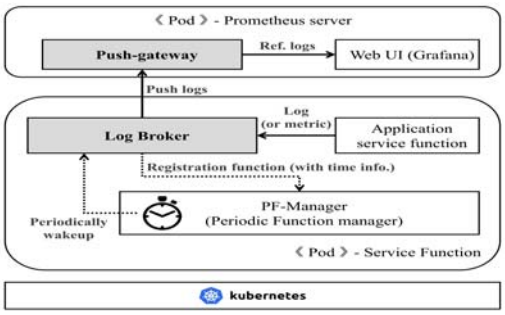


Fig 3. Application structure using the PF-Manager and Log Broker

The PF-Manager consists of a function manager and a timer. The timer checks the periodic operating conditions by accessing the Function Manager at every operation. In this way, the PF-Manager guarantees the periodic operation of registered functions.

The PF-Manager consists of a function manager and a timer as shown in **Fig. 4**. First, the timer is the timer of the operating system used by the Docker container. Function Manager is a data structure that manages the information registrated through the function registration phase. The timer checks the periodic operating conditions by accessing the Function Manager at every operation. In the operation condition checking process, a function that satisfies the condition operates, and a function that does not operate. In this way, the PF-Manager guarantees the periodic operation of registered functions.

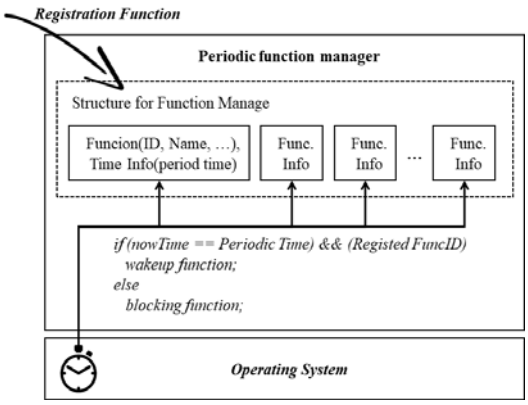


Fig 4. Structure of PF-Manager

4. Experiment

In the experiments of the modules developed through this study, we check the functions of the PF-Manager and the log broker using the open-source FaaS platform. The FaaS platform used in this experiment is OpenFaaS, which is an open source repository that is actively updated with program code. **Fig. 5** is a diagram of the experimental environment.

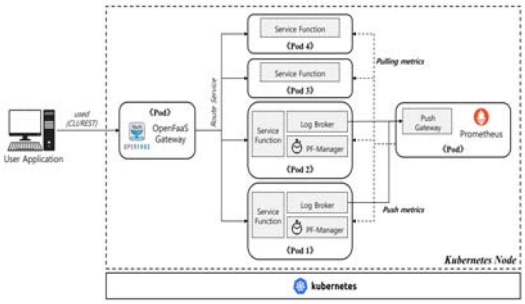


Fig 5. Environment configuration for periodic log broker function

The FaaS applications for the experiment consisted of four Pods (Pod 1-4), which were deployed on the Kubernetes-based OpenFaaS platform. Prometheus collects system level metrics from all FaaS applications. At this time, the FaaS application of Pod 1 and Pod 2 includes PF-Manager and log broker. They send application-level logs to Prometheus' push gateway. For the experiment, the PF-Manager of Pod 1 and Pod 2 executes the log broker each every 1 second and 2 seconds.

Fig. 6 shows the results of function test for the periodic log broker. **Fig. 6-(a)** shows time point when the Prometheus collects system level metrics from each service pod. This results show that the metric collection time is irregular from the respective pods. **Fig. 6-(b)** shows time point when pushed the application level metrics to Prometheus' Push-Gateway by log brokers deployed in FaaS applications Pod 1 and Pod 2. It can be seen that the log broker operates according to the periodic time set by the PF-Manager. It means that the PF-Manager and log broker are designed and implemented well.

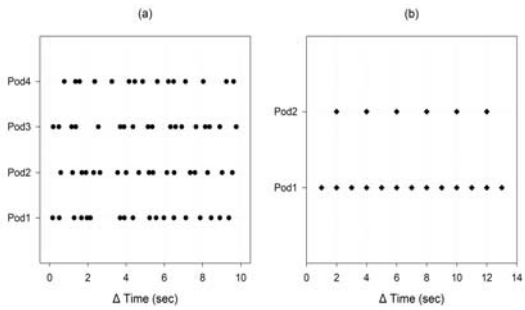


Fig 6. Result of the log broker function

5. Conclusions

In this study, it was confirmed that the developed PF-Manager and the log broker confirmed that they were correctly designed and developed through functional verification experiments. The developed modules confirm that periodic service execution is possible even in cloud-based FaaS applications. It could be hypothesized that future FaaS applications could be used for time-sensitive applications such as edge computing and edge robots. This research will be of great help to researchers studying cloud, FaaS, and real-time systems.

References

[1] R. Kandan, M. F. Khalid, B. I. Ismail, E. M. Goortani, M. N. M. Mydin and O. H. Hoe, "CLOF: A proposed containerized log management orchestration framework," *2017 IEEE Conference on Open Systems (ICOS)*, Miri, pp.13-16, 2017.

[2] Satnam Singh, “Cluster-level Logging of Containers with Containers”, *Queue – Microservices (ACM)*, Vol 14, Issue 3, pp. 1-30, 2016.

KSII The 12th International Conference on Internet (ICONI) 2020.

Design pattern supporting API unit service authorization in FaaS environment : FA Pattern

Dongmin Kim, Kiman Jeon and Jae Gi Son
Department of Human IT Convergence Research Center, KETI
Gyeonggi-do, South Korea
[e-mail: dmkim@keti.re.kr, kmjeon@keti.re.kr, jgson@keti.re.kr]
*Corresponding author: Dongmin Kim

Abstract

FaaS (Function as a Service) is a computing environment that divides an application into a function or container unit, and executes and manages it in a distributed computing or cloud environment. FaaS environments provided by global cloud service companies such as Google, Amazon, and Microsoft provide services centered on FaaS developers and service users, which depend on the service environment provided by companies. FaaS requiring security functions must be secured in all resource units (Function or Container unit that managed in FaaS environment). However, the existing FaaS does not provide a security function in units of functions or containers, and has a service management structure that does not consider users or grouped users, making it difficult to flexibly operate FaaS. Therefore, in this paper, we propose the FA (FaaS Authentication) Pattern, a design pattern that supports authorization functions for all resource and service functions constituting the FaaS environment for flexible operation of FaaS requiring security functions. FA Pattern will increase the usability of the FaaS environment to which enhanced security functions are applied through verification in the open source based FaaS environment that is not dependent on the service environment.

***Keywords*:** API Authentication, Container Authentication, FaaS, FaaS design pattern, Function Authentication

1. Introduction

FaaS (Function as a service) is a computing environment that divides an application service into units that can be called by API unit, and executes and manages it in a distributed computing or cloud environment. FaaS provides a service environment for global cloud service companies such as Google, Amazon, and Microsoft. However, the service structure and types provided by each service vendor are different, and there are limitations in flexible FaaS utilization because they provide dependent services.

Open source-based FaaS computing environments solve the existing problems of commercial computing environments that provide dependent services such as OpenFaaS, Fission, and OpenFx. Operating a FaaS environment using open source must resolve security issues on its own. In the FaaS environment, to solve the security problem, authentication and authorization functions must be provided in every resource unit constituting

This work was supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No.2020-0-00985, Development of Lightweight System Software for Edge Server System Resource Management and Control)

the FaaS environment, but there is no open source-based FaaS environment to solve this problem.

Therefore, in this paper, we propose the FaaS Authentication Pattern (FA Pattern), a software structure to provide authentication and authorization functions for all resources constituting the open source-based FaaS environment. FA Pattern is a design pattern that helps users who want to operate an open source based FaaS environment. This was verified by applying it to OpenFx, and it will be used as a general FaaS-based design pattern as it can be applied to management of edge computing services that require operation and management in service units.

2. Related Works

2.1 OpenFx

OpenFx[1] is an open source-based framework that supports FaaS computing. It supports faster responsiveness than other frameworks using gRPC as a communication interface. **Fig. 1** is the SW structure of OpenFx. This has a structure similar to that of a general FaaS environment, except for the communication interface using gRPC.

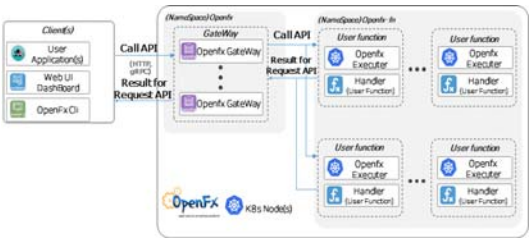


Fig. 1. Structure of FaaS based open source : OpenFx

2.2 FaaS authentication structure based open source

OpenFaaS provides an authentication function for security among well-known open source based FaaS environments. OpenFaaS is to protect FaaS API Gateway when operating OpenFaaS in a public environment. Authentication of OpenFaaS provides an authentication function to access API Gateway by utilizing Basic access authentication function in HTTP communication. However, it provides a partial security function that does not take into account the FaaS environment, and access authentication is not considered for all resources constituting the FaaS environment.

OpenFx provides an authorization function for FaaS service units [1], but the SW structure is not structural and can be used only in a specialized environment, so its generality is low.

3. FaaS Authentication Pattern : FA Pattern

Fig. 2 shows the FA pattern structure that provides security functions for all resources that make up the FaaS environment.

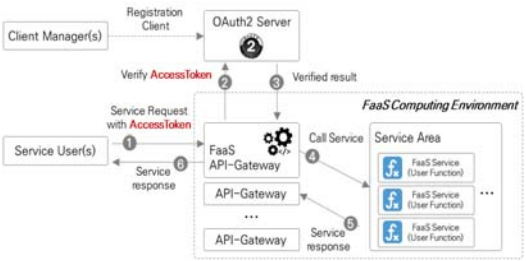


Fig. 2. Structure of FA Pattern

FA Pattern uses OAuth2 protocol to provide security functions of all resources constituting FaaS. OAuth2 is an open standard protocol for delegating access to services with access authority. (It is assumed that the information of the client to be served in the FaaS environment is pre-registered through the client manager.)

When a user in a FaaS environment requests FaaS service to API Gatewayt, the access token information provided from the OAuth2 server is delivered together. API Gateway requests the received token information from the OAuth2 server to verify the access token, executes the request service for a request that allows access, and delivers the result to the user.

4. Experiment and results

The function and performance were verified by applying the FA Pattern proposed in this paper to OpenFx, an open source based FaaS environment. **Fig. 3** shows the OpenFx structure

with FA Pattern applied.

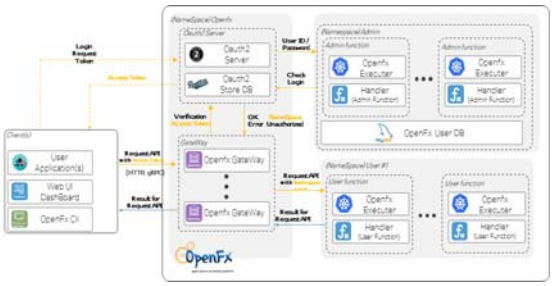


Fig. 3. Structure of FaaS embedded the FA Pattern

It was confirmed that OpenFx with FA Pattern applied correctly provides the access authorization function in the requested service unit (**Fig. 3**).

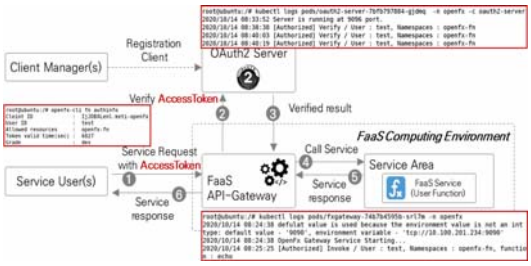


Fig 4. Result of Access token info during call faas service with OAuth2

Also, It is known that the average service response time in an open source based FaaS environment is 180ms. As the average service response time of the Faas environment to which the FA Pattern was applied was confirmed to be 300ms, it was confirmed that the FA Pattern can be used in the FaaS environment without excessive overhead.

5. Conclusions

In this study, the FA Pattern to provide security functions in FaaS environment was proposed, and it was confirmed that there is no problem in terms of function and performance. In addition to the FaaS environment, the proposed function is expected to provide an authorization function in edge computing that manages security functions in service units. This can be applied to all general-purpose environments that provide services in a functional unit in a cloud environment, and based on this, it will be useful for administrators and service developers who manage the service operating environment.

References

[1] Dongmin Kim et al., "An Open Source Serverless Platform by API Call Unit Applications Supported Fast Response : OpenFx", *KIISE Transactions on Computing Practices*, 25(11), pp.543-551, 2019.

[2] Seungho Han et al., “A study on access control technology using oauth2 in serverless framework”, *2020 Summer conference in KINGPC*, pp.278-281, 2020.

A video shortening tool for short-form video creation

Dalwon Jang, Jaewon Lee, Miji Kim, Jai-Won Chung and JongSeol Lee
Communications and media R&D Division
Korea Electronics Technology Institute
[e-mail: dalwon@keti.re.kr, jwlee0121@keti.re.kr, miji0425@keti.re.kr, chjw0915@keti.re.kr, leejs@keti.re.kr]
*Corresponding author: Dalwon Jang

Abstract

This paper presents a video shortening tool which aims to cut the unwanted scene for short-form video service. With the increase of short-form video service market, ordinary users, who do not have a video editing ability, can make a well-decorated video. However, it is not easy for users to cut the video in scene-level. The tool computes a value that represents focus area after scene segmentation, and the value is used as a criterion of scene cutting. In this paper, the structure of the tool and test results with toy example are presented. It will be a part of template-based short-form video making service after more developments.

***Keywords*:** Video Shortening, Video editing, Scene segmentation, Defocus blur detection

1. Introduction

Nowdays, short-form video has been interested, and its markets are also increasing. Worldwide short-form video service TikTok is a front-runner of the market. Recently, a short-form video service named as “blog moment” was opened in Naver. For users to make short-from video with ease, the service provider generally gives a set of template. The user can set their own video, image and text in a template, and finally get a well-decorated short-form video.

In this paper, a video shorting toll for the user-recorded video is introduced. To fix users’ video within a template, it should be edited to follow the fixed time duration and spatial size of the template. Problem of spaial size is easily solved with resizing. When user’s video is longer than the time duration of template, it shoud be shortened. The shortening tool proposed in this paper cuts unfocused segment using defocus area detection[1,2]. This tool does not perform a whole process to adjust time duration, and can be used as pre-processing which first adjusts the duration of users’ video.

2. Structure

Fig. 1 shows the structure and the usage of the video shortening tool. After selecting template, a user uploads his/her own video on short-form video service. Before automatic speed change, unwanted scenes are deleted in the shortening tool.
The input video is first divided depending on the scene[3,4]. Commonly, in recorded video which does not edited, scene segmentation can be interpreted as shot change detection, thus this can be easily done. Detailed explanation is written in [5]. For frames of each segment, defocus blur detection is performed, and we utilize LBP(local

This research is supported by Ministry of Culture, Sports and Tourism and Korea Creative Content Agency(Project Number: R2020040068)

binary pattern)-based method[1]. After defocus blur detection, a gray-scle image is obtained, and the high value in the image means focused area. Here, we assume that the scene not having focused area is unwanted scene, and the scene is deleted with the assumption. For this, a index value, which represents the degree of focusing. is necessary. The index value is decided as the maximum value of Gaussian-blurred image. Example of these processes is shown in **Fig. 2**. Defocus blur detection and index computation are performed every frame, and an average in a scene is used as a criterion. Scenes which have lower average index values than a fixed threshold will be deleted, and the shortened video is obtained. The video is once more processed with speed change in order to exactly adjust the duration time.

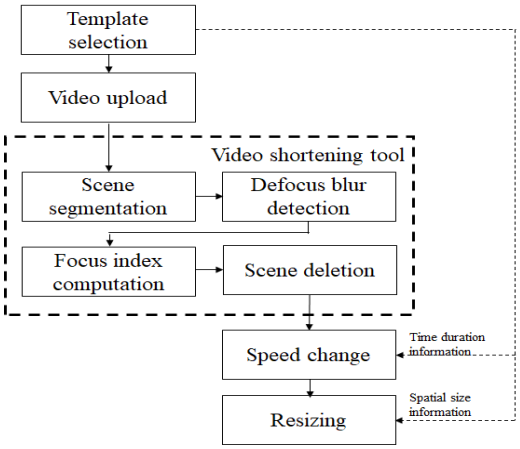


Fig. 1. Structure of the tool

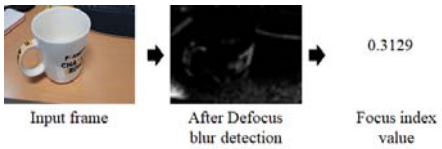


Fig. 2. Example of processes

3. Experimental results

To show the effectiveness of the tool, experiment with a toy example is performed. Input video is 10s length recorded video, and it contains three scenes. First, a cup is recorded, and a mouse is recorded after moving the camera. If the tool operates following its purpose, the moving scene should be deleted. The results are shown in **Fig. 3**. In the figure, representative frames of three scenes are shown, and the graph of index value is shown. From the graph, we can know the moving scene can be decided as being deleted by thresholding the scene average of the value.

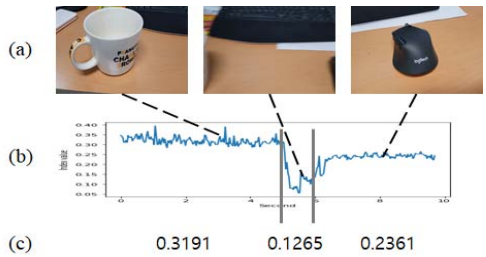


Fig. 3. Results for toy example: (a) Representative f rames of three scenes, (b) Focus index vlaues, an d (c) Averages of scenes

4. Conclusions

In this paper, we presents a video shortening tool which is developed for short-form video service. The tool is based on scene segmentation and defocus blur detection. Using these two, the tool cut the unwanted scene. Developments of the detailed module in the tool is not yet finalized. In future works, we will enhance the tool and present experiments with various video inputs.

References

[1] X. Yi and M. Eramian, “LBP-based segmentation of defocus blur,” IEEE Trans. Image Process., vol. 25, no. 4, pp. 1626–1638, Apr. 2016.

[2] W. Zhao, F. Zhao, D. Wang, and H. Lu., “Defocus blur detection via multi-stream bottom-top-bottom fully convolutional network,” In IEEE Conference on Computer Vision and Pattern Recognition, pages 3080–3088, 2018.

[3] A.Chergui, A.Bekkhoucha, W.Sabbar, “Video scene segmentation using the shot transition detection by local characterization of the points of interest”. SETIT 2012.

[4] C. Huang and H. Wang, “A Novel key-frames selection framework for comprehensive video summarization,” IEEE TCSVT, 2019.

[5] J. Lee, D. Jang, and J.S. Lee, “Automatic Video Splitting System based on Inter-frame Similarities,” ICONI, 2020.

A Priori SNR Estimation with a Weight Adaptation for Speech Enhancement

Jinsoo Seo[1], Jaewon Lee[2], JongSeol Lee[2], and Dalwon Jang[2]
[1] Department of Electrical Engineering, Gangneung-Wonju National University
Gangneung, South Korea
[2] Communications and media R&D Division, Korea Electronics Technology Institute
Seoul, South Korea
[e-mail: jsseo@gwnu.ac.kr, jwlee0121@keti.re.kr, leejs@keti.re.kr dalwon@keti.re.kr]
*Corresponding author: Dalwon Jang

Abstract

A priori signal-to-noise estimation is essential for speech enhancement. This paper proposes an adaptive weight update method for the decision-directed a priori SNR estimation approach to relieve its slow response to abrupt signal change. Experiments on a publicly-available speech dataset show that the proposed weight update is effective in improving the quality of the speech enhancement.

***Keywords*:** Speech enhancement, a priori SNR, decision-directed approach, weight adaptation

1. Introduction

Spectral-domain speech enhancement methods modify noisy speech relying on two key parameters: a priori signal-to-noise ratio (SNR) and a posteriori SNR. Inaccurate SNR estimation causes residual noise or unnatural background sounding (called musical noise). The decision-directed (DD) approach, proposed by Ephraim and Malah [1], has been adopted as a standard method for SNR estimation due to its simplicity and efficiency in reducing musical noise. Analysis by Cappe [2] shows that the a priori SNR estimated by DD approach follows the shape of the a posteriori SNR with a frame delay which leads to slow response to an abrupt speech or noise change (known as the reverberation effect). The slow response of the DD approach is due to the fixed weight factor (which is set close to one amenable for reducing musical noise). To cope with the shortcoming of the DD approach while maintaining its merit, a number of studies has been performed [3,4]. In line with the previous studies, this paper proposes an adaptive weight update method by checking a novelty of the current frame's a posteriori SNR. Only when the novelty of the a posteriori SNR is statistically significant, the weight factor is accordingly changed.

2. Proposed A Priori SNR Estimation

2.1 Speech Enhancement Based on Statistical Models in Spectral Domain

In the uncorrelated additive noise model where the noisy signal $y[n]$ is given by the summation of the clean speech $s[n]$ and the noise $v[n]$, the short-time Fourier transform (STFT) of the noisy signal is expressed by

$$Y_m(k) = S_m(k) + V_m(k) \tag{1}$$

where k and m denote the frequency bin and the frame index respectively. The goal of the speech enhancement in the spectral domain is to

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2020-0-02001)

compute an estimate $\hat{S}_m(k)$ of the clean speech's STFT $S_m(k)$, which is represented by a spectral gain function $G_m(k)$ as follows:

$$\hat{S}_m(k) = G_m(k)Y_m(k) \quad (2)$$

where $G_m(k)$ is determined by a distortion measure and statistical models for speech and noise. The spectral gain $G_m(k)$ is a function of the a priori signal-to-noise ratio (SNR) $\xi_m(k) = \frac{E[S_m^2(k)]}{\lambda_m(k)}$ and the a posteriori SNR $\gamma_m(k) = \frac{E[Y_m^2(k)]}{\lambda_m(k)}$ where $\lambda_m(k)$ is the variance of $V_m(k)$. The estimation of a priori SNR and a posteriori SNR are critical to speech enhancement.

2.2 Proposed Weight Adaptation for a Priori SNR Estimation

The decision-directed a priori SNR estimation [1] is given by

$$\xi_m(k) = a\eta_{m-1}(k) + (1-a)\max[\gamma_m(k) - 1,0] \quad (3)$$

where a is a weighting factor (typically 0.98), and $\eta_{m-1}(k) = \frac{\hat{S}_{m-1}^2(k)}{\lambda_{m-1}(k)}$. A small value of a produces significant amount of musical noise, which support the value of a close to one. However; when a is close to one, it takes a longer time for ξ to reach the final value during signal onset and offset. This paper cope with the problem by making the smoothing constant a to change adaptively. We assume that both the speech and the noise spectral amplitudes follow exponential distributions, whose ratio, $\log(\gamma)$, follows a Logistic distribution with the mean $\max[\log(\eta_{m-1}(k)),0]$. Only when the value of the $\log(\gamma)$ is located at the left or the right tail of the Logistic distribution (lower 0.1 or upper 0.1 quantile), the weighting factor a is changed proportional to its quantile value. The proposed adaptive weight is close to zero only when statistically significant deviation of the γ occurs, which prevents an artificial musical noise.

3. Results and Discussion

Evaluation of the proposed weight adaptation was conducted using the NOIZEUS dataset containing 30 clean IEEE sentences (spoken by 3 male and 3 female speakers). Each clean sentence was corrupted by various real-world noises with the global SNR levels of 0 and 5 dB. A frame size of 20ms with 50% overlap was used. The MMSE spectral gain [1] with minimum statistics noise estimation was used for speech enhancement. The segmental SNR improvement was used as an evaluation measure.

As shown in **Table 1**, the proposed adaptive decision-directed (ADD) estimation outperforms the previous DD approach for all cases in terms of the segmental SNR improvement with the margin of at least 0.15 dB (up to 0.62 dB) depending on the type of noise. The segmental SNR improvement brought by the ADD is explained by its ability to preserve on and offset of the signal and to suppress noise more aggressively during speech pause.

Table 1. Segmental SNR improvement for the NOIZEUS dataset

SNR	0 dB		5 dB	
Noise	DD	ADD	DD	ADD
Airport	1.657	1.914	1.477	1.803
Babble	1.534	1.692	1.447	1.702
Car	2.689	3.300	2.270	2.893
Hall	1.996	2.318	1.843	2.248
Restaurant	1.300	1.457	1.117	1.358
Street	1.911	2.213	1.784	2.200

References

[1] Y. Ephraim and D. Malah, "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator," *IEEE Trans. Acoust. Speech Signal Process.*, Vol 32, No. 6, pp. 1109-1121, 1984.

[2] O. Cappe, "Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor," *IEEE Trans. Speech Audio Process.*, Vol 2, No. 2, pp. 345-349, 1994.

[3] P. C. Yong, S. Nordholm, and H. H. Dam, "Optimization and evaluation of sigmoid function with a priori SNR estimate for real-time speech enhancement," *Speech Communication*, Vol 55, No. 2, pp. 345-349, 2013.

[4] R. Yao, Z. Zeng, and P. Zhu, "A priori SNR estimation and noise estimation for speech enhancement," *EURASIP Journal on Advances in Signal Processing*, No. 101, 2016.