Lecture Notes in Electrical Engineering 536

James J. Park
Doo-Soon Park
Young-Sik Jeong
Yi Pan *Editors*

Advances in Computer Science and Ubiquitous Computing

CSA-CUTE 2018

Springer

Lecture Notes in Electrical Engineering

Volume 536

Series Editors

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Associate Editor (jasmine.dou@springer.com)

India

Aninda Bose, Senior Editor (aninda.bose@springer.com)

Japan

Takeyuki Yonezawa, Editorial Director (takeyuki.yonezawa@springer.com)

South Korea

Smith (Ahram) Chae, Editor (smith.chae@springer.com)

Southeast Asia

Ramesh Nath Premnath, Editor (ramesh.premnath@springer.com)

USA, Canada:

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries:

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** Indexing: The books of this series are submitted to ISI Proceedings, EI-Compendex, SCOPUS, MetaPress, Web of Science and Springerlink ****

More information about this series at http://www.springer.com/series/7818

James J. Park · Doo-Soon Park ·
Young-Sik Jeong · Yi Pan
Editors

Advances in Computer Science and Ubiquitous Computing

CSA-CUTE 2018

Springer

Editors
James J. Park
Department of Computer Science
and Engineering
Seoul National University of Science
and Technology
Seoul, Korea (Republic of)

Young-Sik Jeong
Department of Multimedia Engineering
Dongguk University
Seoul, Korea (Republic of)

Doo-Soon Park
Department of Computer Software
Engineering
Soon Chun Hyang University
Asan, Korea (Republic of)

Yi Pan
Department of Computer Science
Georgia State University
Atlanta, GA, USA

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-13-9340-2 ISBN 978-981-13-9341-9 (eBook)
https://doi.org/10.1007/978-981-13-9341-9

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Message from the CUTE 2018 General Chairs

On behalf of the organizing committees, it is our pleasure to welcome you to the 13th International Conference on Ubiquitous Information Technologies and Applications (CUTE 2018) which will be held in Kuala Lumpur, Malaysia, on December 17–19, 2018.

This conference provides an international forum for the presentation and showcase of recent advances on various aspects of ubiquitous computing. It will reflect the state of the art of the computational methods, involving theory, algorithm, numerical simulation, error and uncertainty analysis, and/or novel application of new processing techniques in engineering, science, and other disciplines related to ubiquitous computing.

The papers included in the proceedings cover the following topics: ubiquitous communication and networking, ubiquitous software technology, ubiquitous systems and applications, ubiquitous security, privacy, and trust. Accepted papers highlight new trends and challenges in the field of ubiquitous computing technologies. We hope you will find these results useful and inspiring for your future research.

We would like to express our sincere thanks to steering committees: James J. Park (SeoulTech, Korea), Doo-Soon Park (Soonchunhyang University, Korea), Young-Sik Jeong (Dongguk University, Korea), Hsiao-Hsi Wang (Providence University, Taiwan), Laurence T. Yang (St. Francis Xavier University, Canada), Hai Jin (Huazhong University of Science and Technology, China), Chan-Hyun Youn (KAIST, Korea), Jianhua Ma (Hosei University, Japan), Mingyi Guo (Shanghai Jiao Tong University, China), Weijia Jia (City University of Hong Kong, Hong Kong). We would also like to express our cordial thanks to the Program Chairs and Program Committee members for their valuable efforts in the review process, which helped us to guarantee the highest quality of the selected papers for the conference.

Finally, we would thank all the authors for their valuable contributions and the other participants of this conference. The conference would not have been possible without their support. Thanks are also due to the many experts who contributed to making the event a success.

Yi Pan
Sanghoon Kim
Luis Javier Garcia Villalba
CUTE 2018 General Chairs

Message from the CUTE 2018 Program Chairs

Welcome to the 13th International Conference on Ubiquitous Information Technologies and Applications (CUTE 2018), which will be held in Kuala Lumpur, Malaysia, on December 17–19, 2018.

The purpose of the CUTE 2018 conference is to promote discussion and interaction among academics, researchers, and professionals in the field of ubiquitous computing technologies. This year the value, breadth, and depth of the CUTE 2018 conference continue to strengthen and grow in importance for both the academic and industrial communities. This strength is evidenced this year by having the highest number of submissions made to the conference.

For CUTE 2018, we received a lot of paper submissions from various countries. Out of these, after a rigorous peer-review process, we accepted only high-quality papers for CUTE 2018 proceeding, published by the Springer. All submitted papers have undergone blind reviews by at least two reviewers from the technical program committee, which consists of leading researchers around the globe. Without their hard work, achieving such a high-quality proceeding would not have been possible. We take this opportunity to thank them for their great support and cooperation.

Finally, we would like to thank all of you for your participation in our conference, and also thank all the authors, reviewers, and organizing committee members. Thank you and enjoy the conference!

Muhammad Khurram Khan
Neil Y. Yen
Yunsick Sung
CUTE 2018 Program Chairs

Organization

Honorary Chair

Seok-Woo Nam (KIPS President)	Comtec Systems Co. Ltd, Korea

Steering Committee

James J. Park (Leading Chair)	SeoulTech, Korea
Doo-Soon Park (Co-chair)	Soonchunhyang University, Korea
Young-Sik Jeong (Co-chair)	Dongguk University, Korea
Hsiao-Hsi Wang	Providence University, Taiwan
Laurence T. Yang	St. Francis Xavier University, Canada
Hai Jin	Huazhong University of Science and Technology, China
Chan-Hyun Youn	KAIST, Korea
Jianhua Ma	Hosei University, Japan
Mingyi Guo	Shanghai Jiao Tong University, China
Weijia Jia	City University of Hong Kong, Hong Kong

General Chairs

Yi Pan	Georgia State University, USA
Sanghoon Kim	Hankyong National University, Korea
Luis Javier Garcia Villalba	Universidad Complutense de Madrid, Spain

Program Chairs

Muhammad Khurram Khan	King Saud University, Kingdom of Saudi Arabia
Neil Y. Yen	The University of Aizu, Japan
Yunsick Sung	Dongguk University, Korea

International Advisory Committee

Witold Pedrycz	University of Alberta, Canada
Seok Cheon Park	Gachon University, Korea
C. S. Raghavendra	University of Southern California, USA
Im-Yeong Lee	Soonchunhyang University, Korea
HeonChang Yu	Korea University, Korea
Hai Jin	Huazhong University of Science and Technology, China
Nammee Moon	Hoseo University, Korea
Byeong-Seok Shin	Inha University, Korea
Dong-Ho Kim	Soongsil University, Korea
Shu-Ching Chen	Florida International University, USA
Keun Ho Ryu	Chungbuk National University, Korea
JaeKwang Lee	Hannam University, Korea
Victor Leung	University of British Columbia, Canada
Yoo-jae Won	Chungnam National University, Korea
Yang Xiao	University of Alabama, USA

Publicity Chairs

Byoungwook Kim	Dongguk University, Korea
Jin Wang	Changsha University of Science and Technology, China
Deok Gyu Lee	Seowon University, Korea
Hyun-Woo Kim	Dongguk University, Korea
Seokhong Min	Mindata Ltd, Korea
Joon-Min Gil	Catholic University of Daegu, Korea
Sung Chul Yu	LG Hitachi Ltd., Korea
Yu-Wei Chan	Providence University, Taiwan
Jaehwa Chung	Korea National Open University, Korea
Jinho Park	Soongsil University, Korea
Hang-Bae Chang	Chung-Ang University, Korea

Program Committee

Bo-Chao Cheng	National Chung Cheng University, Taiwan
Dumitru Roman	SINTEF/University of Oslo, Norway
Imad Saleh	University of Paris 8, France
Jin-Hee Cho	US Army Research Laboratory, USA
Jong-Myon Kim	University of Ulsan, Korea
Kwang Sik Chung	Korea National Open University, Korea
Chen Uei-Ren	Hsiuping University of Science and Technology, Taiwan
Damien Sauveron	University of Limoges, France

Kwangman Ko	Sangji University, Korea
Lai Kuan-Chu	National Taichung University, Taiwan
Loh Woong-Kee	Sungkyul University, Korea
Milan Markovic	Banca Intesa ad Beograd, Serbia
Neungsoo Park	Konkuk University, Korea
Pinaki A Ghosh	Atmiya Institute of Technology and Science, India
Pyung-Soo Kim	Korea Polytechnic University, Korea
Seung-Ho Lim	Hankuk University of Foreign Studies, Korea

Message from the CSA 2018 General Chairs

International Conference on Computer Science and its Applications (CSA 2018) is the 10th event of the series of international scientific conference. This conference takes place in Kuala Lumpur, Malaysia, on December 17–19, 2018. CSA 2018 will be the most comprehensive conference focused on the various aspects of advances in computer science and its applications. CSA 2018 will provide an opportunity for academic and industry professionals to discuss the latest issues and progress in the area of CSA. In addition, the conference will publish high-quality papers which are closely related to the various theories and practical applications in CSA. Furthermore, we expect that the conference and its publications will be a trigger for further related research and technology improvements in this important subject. CSA 2018 is the next event in a series of highly successful International Conference on Computer Science and its Applications, previously held as CSA 2017 (9th Edition: Taichung, Taiwan), CSA 2016 (8th Edition: Bangkok, Thailand, 2016), CSA 2015 (7th Edition: Cebu, December 2015), CSA 2014 (6th Edition: Guam, December 2014), CSA 2013 (5th Edition: Danang, December 2013), CSA 2012 (4th Edition: Jeju, November 2012), CSA 2011 (3rd Edition: Jeju, December 2011), CSA 2009 (2nd Edition: Jeju, December 2009), and CSA 2008 (1st Edition: Australia, October 2008).

The papers included in the proceedings cover the following topics: mobile and ubiquitous computing, dependable, reliable and autonomic computing, security and trust management, multimedia systems and services, networking and communications, database and data mining, game and software engineering, grid and scalable computing, embedded system and software, artificial intelligence, distributed and parallel algorithms, Web and Internet computing, and IT policy and business management.

Accepted and presented papers highlight new trends and challenges of computer science and its applications. The presenters showed how new research could lead to novel and innovative applications. We hope you will find these results useful and inspiring for your future research. We would like to express our sincere thanks to Steering Chairs: James J. (Jong Hyuk) Park (SeoulTech, Korea), Yi Pan (Georgia State University, USA), Han-Chieh Chao (National Ilan University, Taiwan),

Young-Sik Jeong (Dongguk University, Korea), Vincenzo Loia (University of Salerno, Italy).

Our special thanks go to the Program Chairs: Arun Kumar Sangaiah (VIT University, India), Mu-Yen Chen (National Taichung University of Science and Technology, Taiwan), Houcine Hassan (Universitat Politècnica de València, Spain), all Program Committee members, and all the additional reviewers for their valuable efforts in the review process, which helped us to guarantee the highest quality of the selected papers for the conference.

We cordially thank all the authors for their valuable contributions and the other participants of this conference. The conference would not have been possible without their support. Thanks are also due to the many experts who contributed to making the event a success.

Kim-Kwang Raymond Choo
Victor Leung
Jungho Kang
CSA 2018 General Chairs

Message from the CSA 2018 Program Chairs

Welcome to the 10th International Conference on Computer Science and its Applications (CSA 2018) which will be held in Kuala Lumpur, Malaysia, on December 17–19, 2018. CSA 2018 will be the most comprehensive conference focused on the various aspects of advances in computer science and its applications.

CSA 2018 provides an opportunity for academic and industry professionals to discuss the latest issues and progress in the area of computer science. In addition, the conference contains high-quality papers which are closely related to the various theories and practical applications in computer science. Furthermore, we expect that the conference and its publications will be a trigger for further related research and technology improvements in this important subject. CSA 2018 is the next event in a series of highly successful International Conference on Computer Science and its Applications, previously held as CSA 2017 (9th Edition: Taichung, Taiwan), CSA 2016 (8th Edition: Bangkok, Thailand, 2016), CSA 2015 (7th Edition: Cebu, December 2015), CSA 2014 (6th Edition: Guam, December 2014), CSA 2013 (5th Edition: Danang, December 2013), CSA 2012 (4th Edition: Jeju, November 2012), CSA 2011 (3rd Edition: Jeju, December 2011), CSA 2009 (2nd Edition: Jeju, December 2009), and CSA 2008 (1st Edition: Australia, October 2008).

CSA 2018 contains high-quality research papers submitted by researchers from all over the world. Each submitted paper was peer-reviewed by reviewers who are experts in the subject area of the paper. Based on the review results, the Program Committee accepted papers.

For organizing an international conference, the support and help of many people are needed. First, we would like to thank all authors for submitting their papers. We also appreciate the support from Program Committee members and reviewers who carried out the most difficult work of carefully evaluating the submitted papers.

We would like to give my special thanks to Prof. James J. (Jong Hyuk) Park, Prof. Yi Pan, Prof. Han-Chieh Chao, Prof. Young-Sik Jeong, and Prof. Vincenzo Loia as the Steering Committee Chairs of CSA for their strong encouragement and

guidance to organize the symposium. We would like to thank CSA 2018 General Chairs: Prof. Kim-Kwang Raymond Choo, Prof. Victor Leung, Prof. Jungho Kang. We would like to express special thanks to committee members for their timely unlimited support.

Arun Kumar Sangaiah
Mu-Yen Chen
Houcine Hassan
CSA 2018 Program Chairs

Organization

Honorary Chair

Doo-Soon Park	Soonchunhyang University, Korea

Steering Committee

James J. Park (Leading Chair)	SeoulTech, Korea
Yi Pan	Georgia State University, USA
Han-Chieh Chao	National Ilan University, Taiwan
Young-Sik Jeong	Dongguk University, Korea
Vincenzo Loia	University of Salerno, Italy

General Chairs

Kim-Kwang Raymond Choo	The University of Texas at San Antonio, USA
Victor Leung	The University of British Columbia, Canada
Jungho Kang	Baewha Women's University, Korea

Program Chairs

Arun Kumar Sangaiah	VIT University, India
Mu-Yen Chen	National Taichung University of Science and Technology, Taiwan
Houcine Hassan	Universitat Politècnica de València, Spain

International Advisory Committee

Mo-Yuen Chow	North Carolina State University, USA
Simon James Fong	University of Macau, Macau, China
Shu-Ching Chen	Florida International University, USA
Mohammad S. Obaidat	Monmouth University, USA

Enrique Herrera-Viedma	University of Granada, Spain
Sherali Zeadally	University of Kentucky, USA
Jordi Mongay Batalla	National Institute of Telecommunications, Poland
Wanlei Zhou	Deakin University, Australia
Sethuraman Panchanathan	Arizona State University, USA
Yueh-Min Huang	National Cheng Kung University, Taiwan

Publicity Chairs

Ching-Hsien Hsu	Chung Hua University, Taiwan
Ka Lok Man	Xi'an Jiaotong-Liverpool University, China
Wei Song	North China University of Technology, China
Deok Gyu Lee	Seowon University, Korea
Fei Hao	Shaanxi Normal University, China
Neil Y. Yen	University of Aizu, Japan

Program Committee

Andrew Kusiak	The University of Iowa, USA
Cho-Chin Lin	National Ilan University, Taiwan
Eunyoung Lee	Dongduk University, Korea
Jerzy Respondek	Silesian University of Technology, Poland
Kuei-Ping Shih	Tamkang University, Taiwan
Listanti Marco	DIET, Roma, Italy
Paprzycki Marcin	Polish Academy of Sciences, Poland
Schulz Frank	SAP Research, Germany
Jaesoo Yoo	Hanbuk University, Korea
Sungsuk Kim	Sun Moon University, Korea
Watanobe Yutaka	University of Aizu, Japan
Haiduke Sarafian	The Pennsylvania State University, USA
Jung Hanmin	KISTI, Korea
Liu Chuan-Ming	National Taipei University of Technology, Taipei
Maiga Chang	Athabasca University, Canada
Morales M. Dominguez	University of Seville, Spain
Nader F. Mir	San Jose State University, USA
Nelson Passos	Midwestern State University, USA
SeongHo Lim	HUFS, Korea
Somchai Chatvichienchai	University of Nagasaki, Japan
Genge Bela	University of Medicine, Pharmacy, Science and Technology of Targu Mures, Romania
Chia-Hung Yeh	National Sun Yat-sen University, Taiwan
El-Sayed El-Alfy	King Fahd University of Petroleum and Minerals, Saudi Arabia
M. Dominguez Morales	University of Seville, Spain
Qian Yu	University of Regina, Canada

Hoon Choi	Chungnam National University, Korea
Yue-Shan Chang	National Taipei University, Taipei
Alexey Rodionov	Institute of Computational Mathematics and Mathematical Geophysics, Russia
Clement Leung	Hong Kong Baptist University, Hong Kong
Dorairaj Prabu	NetApp, USA
Hong Tzung-Pei	National University of Kaohsiung, Taiwan
Oualkadi Ahmed EL	Abdelmalek Essaadi University, Morocco
Parimala Thulasiraman	University of Manitoba, Canada
Radjef Mohammed Said	University of Bejaia, Algeria
Valev Ventzeslav	Bulgarian Academy of Sciences, Bulgaria

Deep Learning Platform for B5G Mobile Network (Invited Speaker)

Han-Chieh Chao

President National Dong Hwa University, Taiwan

Abstract. The 3G and 4G mobile communications had been developed for many years. The 5G mobile communication is scheduled to be launched in 2020. In the future, a wireless network is of various sizes of cells and different types of communication technologies, forming a special architecture of heterogeneous networks (HetNet). Under the complex network architecture, interference and handover problems are critical challenges in access network. How to efficiently manage small cells and to choose an adequate access mechanism for the better quality of service is a vital research issue. Traditional network architecture can no longer support existing network requirements. It is necessary to develop a novel network architecture. Therefore, this keynote speech will share a solution of deep learning-based B5G mobile network which can enhance and improve communication performance through combing some specific technologies, e.g., deep learning, fog computing, cloud computing, cloud radio access network (C-RAN), and fog radio access network (F-RAN).

Biography:
Han-Chieh Chao received his M.S. and Ph.D. degrees in Electrical Engineering from Purdue University, West Lafayette, Indiana, in 1989 and 1993, respectively.

He is currently Professor with the Department of Electrical Engineering, National Dong Hwa University, where he also serves as President. He is also with the Department of Computer Science and Information Engineering and the Department of Electronic Engineering, National Ilan University, Taiwan; College of Mathematics and Computer Science, Wuhan Polytechnic University, Wuhan, China; and Fujian University of Technology, Fuzhou, China. He was Director of the Computer Center for Ministry of Education Taiwan from September 2008 to July 2010. His research interests include IPv6, cross-layer design, cloud computing, IoT, and 5G mobile networks. He has authored or co-authored 4 books and has published about 400 refereed professional research papers. He has completed more than 150 MSEE thesis students and 11 Ph.D. students. He has been invited frequently to give talks at national and international conferences and research organizations. He serves as Editor-in-Chief for the Institution of Engineering and Technology Networks, the Journal of Internet Technology, the International Journal of Internet Protocol Technology, and the International Journal of Ad Hoc and Ubiquitous Computing. He is Fellow of IET (IEE) and Chartered Fellow of the British Computer Society. Due to his contribution of suburban ICT education, he has been awarded the US President's Lifetime Achievement Award and International Albert Schweitzer Foundation Human Contribution Award in 2016.

Contents

Generation of Similar Traffic Using GAN for Resolving Data Imbalance

Woo Ho Lee(✉), Chae Sang Lim, and Bong Nam Noh

Interdisciplinary Program of Information Security,
Chonnam National University, Gwangju, South Korea
leeouho@naver.com

Abstract. Recently, as the practical application of deep learning has become possible, research on the problems pertaining to intrusion detection has increased.

However, it is difficult to detect a small number of attack traffic when the real network is connected to produce an imbalance between the attack traffic class data and the normal traffic data necessary for learning. In this study, we propose a method to improve the accuracy of attack traffic data detection by creating similar attack traffic, using a Generative Adversarial Network (GAN) algorithm of deep learning. The proposed method generates similar attack traffic for NSL–KDD, ISCX 2012, and USTC_TFC 2016 datasets, which are well-known intrusion detection learning data sets. Experiments have shown that the data imbalance in each data set can improve classification accuracy by 10–12%, owing to the degradation problem.

Keywords: Deep learning · Intrusion detection · Security · Generative Adversarial Network

1 Introduction

Recently, studies on deep learning have been conducted for various purposes in the fields of big data, cloud computing, malware detection, and traffic management. Among them, intrusion detection is one of the fields that is most actively researched by using deep learning. Currently, studies on intrusion detection are being conducted using network traffic datasets that are available to the public. In particular, NSL–KDD [1], ISCX2012 [2], and USTC-TFC2016 [3] datasets are often used as training data. However, datasets used for deep learning have insufficient attack traffic information compared to normal traffic. This can cause an imbalanced data problem in the learning processes that use deep learning. The imbalanced data problem refers to imbalanced ratios of data classes where the accuracy of the class with a large amount of data is high, and the accuracy of the class with a small amount of data is low, thus increasing the difficulty of determining performance. To solve the imbalanced data problem, this study proposes an oversampling method that uses the GAN algorithm to create similar attack traffic in a dataset for learning purposes.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 1–7, 2020.
https://doi.org/10.1007/978-981-13-9341-9_1

The structure of this paper is as follows:

- Section 2 analyzes various datasets and outlines studies on intrusion detection, using deep learning with datasets.
- Section 3 describes the structure and data preprocessing of the proposed method.
- Section 4 describes the details of the experiment.
- Section 5 presents the conclusion and future research subjects.

2 Related Work

2.1 Analysis of Datasets

The datasets generally used on intrusion detection studies using deep learning were analyzed. As shown in Table 1, the NSL–KDD and USTC-TFC 2016 datasets have insufficient attack traffic compared to normal traffic. Furthermore, the DDOS attack traffic accounts for more than 50% of all attack traffic, causing the imbalanced data problem. In the case of the ISCX 2012 dataset, the ratio of the attack traffic that can be used for training to the normal traffic ranges from 1:515 at the minimum to 1:1220 at the maximum, causing the imbalanced data problem.

Table 1. Analysis of NSL–KDD, ISCX 2012, and USTC-TFC 2016 datasets

Whole dataset								
ISCX 2012 Dataset			NSL–KDD Dataset			USTC-TFC 2016		
Traffic		Count	Traffic		Count	Traffic		Count
Normal		41,480	Normal		80,767	Normal		345,407
Bot	Neris	8,039	Ab-normal	Probe	5,356	Ab-normal	Cridex	406,633
	Rbot	6,073		DoS	223,488		Geodo	
	Virut	18,914		U2R	228		Htbot	
	Menti	217		R2L	1,376		Miuref	
	Sogou	34					Neris	
	Mutio	2,013					Nsis-ay	
	Nsis-ay	4,395					virut	
	Total	39,685		Total	230,448		Zeus	
Total		81,165	Total		311,215	Total		752,040

2.2 Deep Learning-Based IDS-Related Research

Recently, studies on intrusion detection using deep learning have diversified and increased rapidly. Yao et al. [4] proposed a classification method using k-means, KNN, decision tree, and random forest algorithms based on the NSL-KDD dataset and a classification method for attack traffic. In this study, the attack detection rate was 95.4% for normal traffic, 93.9% for DOS attacks, 56.1% for probe attacks, and 77.2% for U2R and R2L attacks. The low detection rate of the probe attacks at 56.1% confirmed that the smaller the number of attack data is, the lower the detection rate becomes. Rathore et al. [5] evaluated datasets for real-time processing. They applied random forest tree, conjunctive rule, SVM, and naive Bayes procedures to DARPA, KDD99, and NSL–KDD datasets. This study had a limitation because the accuracy was 99.9% in the experimental environment; however, it was not applicable to an actual environment. Tang et al. [6] used a deep neural network with the NSL–KDD dataset to apply an intrusion detection system (IDS) in a software-defined networking environment. They selected six features among 41. The experiment results showed a loss rate of 7.4%, and an accuracy of 91.7%. They also described a classification method using all 41 features.

Mylavarapu et al. [7] conducted a study on real-time detection using the ISCX2012 dataset. For real-time processing, they applied the multilayer perceptron neural networks based on Storm, and classified traffic at 89% accuracy. Yu et al. [8] proposed a session-based network intrusion detection method using CTU013 and ISCX 2012. This method used the stacked denoising autoencoder and showed an accuracy of 98.11% in the experimental environment. However, studies on datasets in an actual environment are insufficient. Wang et al. [3] suggested a method of applying the convolutional neural network (CNN) model with only several hundred bytes of session traffic, which detects malware traffic in the early stages. This method is characterized by independence from protocols because it classifies images using the CNN algorithm. Research on intrusion detection using deep learning is being conducted to improve the accuracy of intrusion detection by extracting the characteristics of attack data, using undersampling and few-shot methods [9]. Most studies on intrusion detection use the undersampling method to increase the layers of the neural network because it is effective for detecting a few classes [10]. However, the undersampling method shows a low accuracy in an actual environment, although its accuracy in an experimental environment is high. In an actual environment, the imbalanced data problem becomes worse compared to the learning data, and the detection rate decreases [11].

Therefore, in this study, in order to apply the oversampling method as a solution to the imbalanced data problem, a similar traffic generation method is proposed to increase the training data of the attack class traffic using a GAN.

3 Proposed Similar Traffic Creation System Using GAN

3.1 Network Feature Extraction

In this thesis, 15 data set characteristics, such as Duration, Header Length, IPversion, Protocol, Flag, and Session were extracted and correlated to detect malware using

network information. Through correlation analysis, six Features of Duration, TCP Header, Port, Session Data, and Flag were selected and applied to a GAN Algorithm. To adjust the size of the data, the hash value is applied and quantified. This allowed the creation of images of the same size as when they were created. The figure below (Fig. 1) shows the CSV file and the correlation analysis for each Feature.

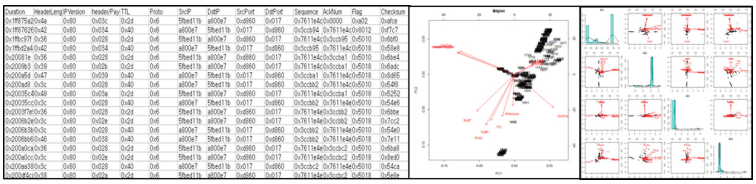


Fig. 1. System architecture for similar traffic generation

3.2 GAN Using Similar Traffic Algorithm

DCGAN is applied to generate network traffic efficiently [12]. Discriminators used in CNN use the usual stride convolution. The generator will produce an image that is as closely as possible similar to the input value. In the case of network traffic, we use Fractionally Stride convolution to increase the size of the Feature-map in case of small images. Discriminator is a Stride Convolution that determines the step size of the kernel when learning an image. It reduces the number of unnecessary parameters and selects only important features when using a normal pooling layer, but it has the disadvantage of losing image position information. In the case of network traffic, we used Stride Convolution because the area is divided and location information is important (Fig. 2).

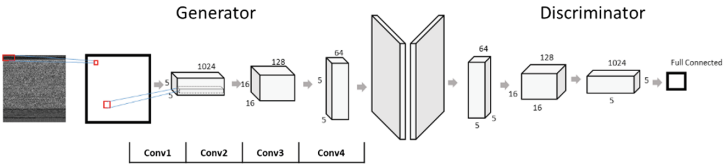


Fig. 2. Illustration of similar traffic created based on a GAN

Algorithm 1 creates images using GAN. From the attack traffic generated through the data-preprocessing step, the GAN algorithm is used to create similar traffic for each class. The criteria for generating similar traffic were between 0.98 and 1 for the loss function of the Generator and between 0 and 1 for the loss function of the Discriminator. The generated similar traffic that meets these criteria was updated to the training data by class. A similar traffic-create criterion generates similar traffic for class loss of 40% of total attack traffic. The created data were only used as training data. If the created similar data would be used for validation data, an overfitting problem can be generated. Thus, it was only used as training data for the test (Fig. 3).

Algorithm 1 GAN Create Algorithm

```
DA= Dataset (Attack Traffic)
DS = Similar Traffic Class
DC = Similar Traffic Create
DT = DA + DC
g = Generate loss
d = Discriminator loss
Function-GAN(DA)
FOR counter i = 0 to Ds length DO
  IF (DA * 40 /100) > (DA * DS[counter i] /
100) THEN
    FOR counter j = 0 to file length DO
      IF d>0.98 && d < 1 THEN
        IF g >0 && g<0.1 THEN
          DT = DA+DC
        ENDIF
      ENDIF
    ENDFOR
  ENDIF
ENDFOR
RETURN with DT
```

Fig. 3. GAN algorithm used for image creation

4 Details of Experiment

4.1 Detection Experiment Using Similar Traffic Dataset

An experiment was conducted to examine the effect of normal/attack traffic created with the GAN Generator on the classification. In this experiment, mirai, virut, smoke_bot, menti, and zeus, which have small amounts of attack traffic, were used.

In the first experiment, the accuracy was analyzed depending on the learning volume using the datasets. The accuracy was measured up to 5,000 epochs under the same conditions for attack traffic of the learning data as before the experiment using GAN.

When the validation test was performed at more than 4,000 epochs, the accuracy was 90–92%, thus improving by approximately 10–12% compared to before the creation of the similar traffic. Furthermore, the learning accuracy for the test data also increased by approximately 7–12%. Therefore, the accuracy improved as the learning volume increased with the similar traffic creation method of the GAN algorithm, and the detection performance was improved by resolving the imbalanced data problem of the datasets (Fig. 4).

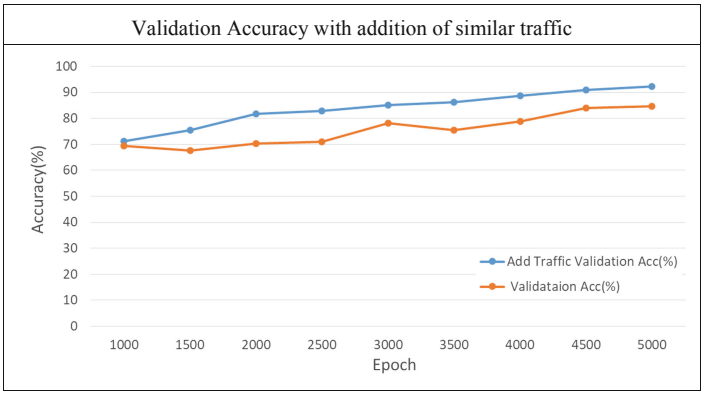


Fig. 4. Measuring and comparing epoch reference accuracies.

In the second experiment, experiments were performed for each data set using the CNN algorithm in a real-time environment.

Table 2 shows the experimental results for classification of attack traffic for NSL-KDD, USTC-TFC 2016, and ISCX 2012 datasets using CNN after adding the data created using GAN to the existing datasets. The classification accuracy of attack traffic in the datasets generally improved. In particular, the dataset validation accuracy increased by 12% and the test accuracy increased by 10%. These results prove that the method of continuously increasing the attack traffic using GAN can be an effective detection method for increasing accuracy in research environment cases with a small amount of attack traffic.

Table 2. Comparisons of attack traffic classification experiments using CNN

Performance	Before Accuracy	After Accuracy	10%		20%		30%	
Dataset			Precision	Recall	Precision	Recall	Precision	Recall
USTC-TFC	83–87	91–95	83.4	83.8	86.2	86.7	91.1	91.6
ISCX-2012	76–81	89–91	82.1	83	85.6	85.3	88.2	89.4
NSL–KDD	75–80	85–93	80.3	80.1	82.1	82.5	84.1	92.1

5 Conclusion

This study investigated a method to improve detection accuracy by creating similar traffic of GAN and increasing the ratio of attack traffic. The method presented in this paper contributed to solving the problem of performance degradation due to data imbalance in the data set using GAN. The existing datasets have an imbalanced data problem, which impedes the detection of attacks or increased learning rate. As a solution to this problem, we added learning data to similar traffic generated by the GAN

algorithm and improved the accuracy in attack detection experiments. In order to create similar traffic, between 0.98 and 1 more images of the Generator was added to the training data. The experiment results showed that the increase of the created similar traffic improved the learning rate by 9–12%, and the detection accuracy for the attack traffic by approximately 13%. Furthermore, we were able to improve classification accuracy according to attack characteristics. In future, classification to determine the type of attack traffic and transformation of the generated images into text form using GAN will be researched.

References

1. Tavallaee, M., Bagheri, E., Lu, W., Ghorbani, A.: A detailed analysis of the KDD CUP 99 data set. Submitted to Second IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA) (2009)
2. Shiravi, A., Shiravi, H., Tavallaee, M., Ghorbani, A.A.: Toward developing a systematic approach to generate benchmark datasets for intrusion detection. Comput. Secur. **31**(3), 357–374 (2012). ISSN 0167-4048. https://doi.org/10.1016/j.cose.2011.12.012
3. Wang, W., et al.: Malware traffic classification using convolutional neural network for representation learning. In: 2017 International Conference on Information Networking (ICOIN). IEEE (2017)
4. Yao, H.P., Liu, Y.Q., Fang, C.: An abnormal network traffic detection algorithm based on big data analysis. Int. J. Comput. Commun. Control **11**(4), 720–733 (2016)
5. Rathore, M.M., Ahmad, A., Paul, A.: Real time intrusion detection system for ultra-high-speed big data environments. J. Supercomput. **72**(9), 3489–3510 (2016)
6. Tang, T.A., et al.: Deep learning approach for network intrusion detection in software defined networking. In: 2016 International Conference on Wireless Networks and Mobile Communications (WINCOM). IEEE (2016)
7. Mylavarapu, G., Thomas, J., Ashwin Kumar, T.K.: Real-time hybrid intrusion detection system using apache storm. In: High Performance Computing and Communications (HPCC), 2015 IEEE 7th International Symposium on Cyberspace Safety and Security (CSS), 2015 IEEE 12th International Conference on Embedded Software and Systems (ICESS), 2015 IEEE 17th International Conference on. IEEE (2015)
8. Yu, Y., Long, J., Cai, Z.: Session-based network intrusion detection using a deep learning architecture. In: Modeling Decisions for Artificial Intelligence. Springer, Cham (2017)
9. He, H., Garcia, E.A.: Learning from imbalanced data. IEEE Trans. Knowl. Data Eng. **21**(9), 1263–1284 (2009)
10. Tavallaee, M., et al.: A detailed analysis of the KDD CUP 99 data set. In: IEEE Symposium on Computational Intelligence for Security and Defense Applications, CISDA 2009. IEEE (2009)
11. Hariharan, B., Girshick, R.: Low-shot visual recognition by shrinking and hallucinating features. In: Proceedings of IEEE International Conference on Computer Vision (ICCV), Venice, Italy (2017)
12. DCGAN: Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434 (2015)

Mid-Level Feature Extractor for Transfer Learning to Small-Scale Dataset of Medical Images

Dong-ho Lee, Yeon Lee, and Byeong-seok Shin(✉)

Department of Computer Engineering, INHA University, Incheon, South Korea
dongholabl3@gmail.com, {leeyeon, bsshin}@inha.ac.kr

Abstract. In fine-tuning-based transfer learning, the size of the dataset may affect the learning accuracy. When a dataset scale is small, fine-tuning-based transfer learning methods use high computing costs, similar to a large-scale dataset. we propose a mid-level feature extractor that only retrains the mid-level convolutional layers, resulting in increased efficiency and reduced computing costs. This mid-level feature extractor is likely to provide an effective alternative in training a small-scale medical image dataset. The performance of the mid-level feature extractor is compared with performance of low- and high-level feature extractors, as well as the fine-tuning method. The mid-level feature extractor takes shorter time to converge than other methods, and it shows good accuracy, obtaining an area under the ROC curve (AUC) of 0.87 in untrained test dataset that is very different from training dataset.

Keywords: Transfer learning · Medical images · Convolutional neural networks · Machine learning

1 Introduction

Convolutional neural networks (CNN) [1] have shown outstanding performance in body detection and disease classification. Several studies [2–4] have demonstrated that the pre-trained model in ImageNet can be transferred to medical images. CheXNet [4] trained pneumonia pathology with better accuracy than doctors could produce using transfer learning. In addition, researchers [2] applied AlexNet [5] within ImageNet to chest x-rays. As a result, they demonstrated that it is possible to transfer non-medical datasets to medical imaging datasets. Other researchers [3] showed that AlexNet and GoogLeNet [6] learned in ImageNet can be used as pre-trained models for transfer learning to CT datasets.

Transfer learning is a means of training new datasets that are limited in size. By using the pre-trained model that trained with sufficiently large-scale datasets during the initialization process of features, small-scale datasets can be trained more efficiently than they could be from the scratch [7]. However, recent CNN models have very deep structures with more than hundreds of layers, so it is not easy to retrain these models with small-scale datasets [7].

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 8–13, 2020.
https://doi.org/10.1007/978-981-13-9341-9_2

Generally, it is difficult to collect large-scale medical imaging datasets due to patient privacy or ethics-related concerns. In the case of rare diseases, small-scale datasets are inevitable. Furthermore, the networks for medical images are applied locally (e.g., the local community). In other words, it is sometimes unnecessary to collect large-scale datasets such as those collected by ImageNet [5]. Furthermore, it is often the case that the necessary infrastructure for collecting medical images is not well established. Therefore, it is necessary to consider an effective deep-learning method for the training of small-scale medical imaging datasets.

Fine-tuning is a method that retrains all layers and classifiers within a network. It is a method that is typically used in datasets from various fields within transfer learning. [2, 3] also have implemented transfer learning through fine-tuning. However, the possibility of overfitting is very high when fine-tuning the pre-trained model to a small-scale dataset. To address this limitation, we utilized feature extraction [8]. The feature extractor method for transfer learning can be thought of in two ways: (1) to prevent updates of all convolutional layers, to only extract features, and then to only train the last fully-connected layer; (2) to divide the convolutional layer into levels and then to train each level selectively. The convolutional layer can be divided into a low level, middle level, and high level, and deeper level extraction displays more complex features closer to the shape of the object [9]. Because each level layer has different features, it is necessary to compare which level is the most effective at retraining small-scale medical imaging datasets.

We propose that the mid-level feature extractor, which retrains features such as the texture of the image and the part of the object, is the most effective model. We compared the transfer learning model, implemented with fine-tuning, and the feature extractor model of each level to the proposed method. As a result, we confirmed that our model has the ability to conduct efficient, small-scale medical image analysis.

2 Mid-Level Feature Extraction for Transfer Learning

The CNN model effectively extracts mid-level image representation [7]. However, as the parameters in training this type of network are very large, learning from an insufficient number of small-scale datasets can lead to overfitting [7]. This problem can be similarly applied to fine-tuning, which retrains all layers. Medical images are limited in their dataset amounts; furthermore, their features are completely different than real-life images, such as are present in ImageNet. As a result, we determined to explore a method that extracts mid-level image features effectively and prevents the overfitting of small-scale datasets.

First, we considered utilizing the construct transfer learning network with the feature extractor method to prevent overfitting by reducing the number of parameters to retrain. Second, we selected which convolutional layer we planned to retrain and update. The low-level layers show the edge and color of the image, the mid-level layers represent the texture of the image or the specific parts of the objects, and the high-level layers show the larger part of the objects or the entire objects [2]. Therefore, we reasoned that the mid-level layers can represent both common and more class-specific features. As a result, we determined to use the mid-level feature extractor method.

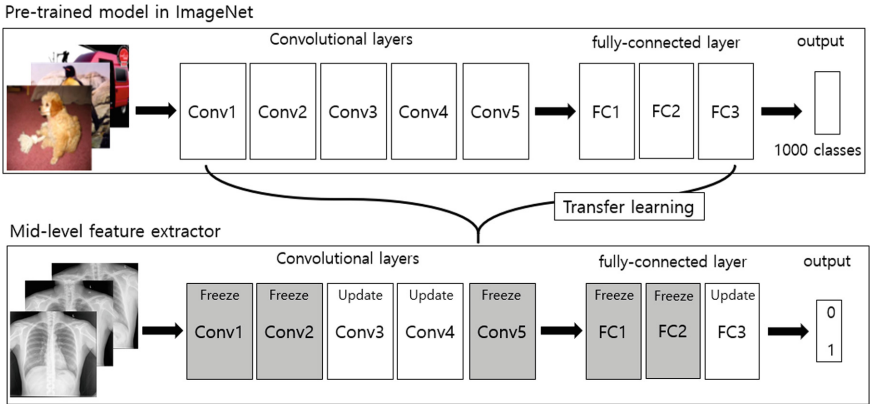


Fig. 1. Mid-level feature extractor structure. Our network is trained on ImageNet. We only train mid-level layers (Conv3, Conv4) and last fully-connected layer.

Figure 1 shows the entire network structure of the mid-level feature extractor. The network is based on using AlexNet in ImageNet, which has a simple structure that allows for the comparison of the effects of training the mid-level convolutional layer. According to [9], AlexNet's conv3 and conv4 refer to textures, such as a mesh pattern, and more complex and class-specific features of the objects in the image. We only trained these two layers. That is, we choose conv3 and conv4 as the mid-level layers. We updated our network by retraining only the mid-level layers, which are conv3, conv4, and the last fully connected layer. Conv1 and conv2, which are low-level layers, and conv5, which is a high-level layer, were frozen, and they maintained the parameters of AlexNet within ImageNet. In addition, the classifiers fc1 and fc2 were frozen, and only fc3 was updated, assuming the possibility of similar experiments in more restrictive situations.

3 Experimental Methods and Results

Our dataset included frontal chest x-ray image data, labeled pulmonary tuberculosis (TB) or non-TB from [10]. There were two datasets, which are Shenzhen dataset from Shenzhen No. 3, People's Hospital (Guangdong providence, China) and Montgomery dataset from Department of Health and Human Services of Montgomery County (MD, USA). While Shenzhen dataset consisted of 336 TB and 326 non-TB, Montgomery dataset consisted of 103 TB and 296 non-TB. Also, Shenzhen dataset was used as a training and validation set, and Montgomery dataset was only used as a test set to examine the possibility of overfitting because the features of Montgomery dataset is very different from Shenzhen dataset. Based on the ImageNet, all input images in the dataset were converted to a size of 224 $\times$ 224 and were normalized.

Unlike ImageNet, the output of our model was one class of TB. Therefore, the output of the last fully connected layer was adjusted to one-dimension. We optimized

the binary cross-entropy (BCE) loss by using the Adam optimizer. Our initial learning rate was 0.001.

Table 1. 9 cases of experiment.

Case	Conv1	Conv2	Conv3	Conv4	Conv5
Mid-level FE	freeze	freeze	update	update	freeze
Conv3-FE	freeze	freeze	update	freeze	freeze
Conv4-FE	freeze	freeze	freeze	update	freeze
Fine-tuning	update	update	update	update	update
Low-level FE	update	update	freeze	freeze	freeze
Conv1-FE	update	freeze	freeze	freeze	freeze
Conv2-FE	freeze	update	freeze	freeze	freeze
High-level FE	freeze	freeze	freeze	freeze	update
H-M level FE	freeze	freeze	update	update	update

The experiment was divided into 9 cases, depending on whether or not each convolutional layer was frozen. Table 1 shows which layers updated or did not update during the training.

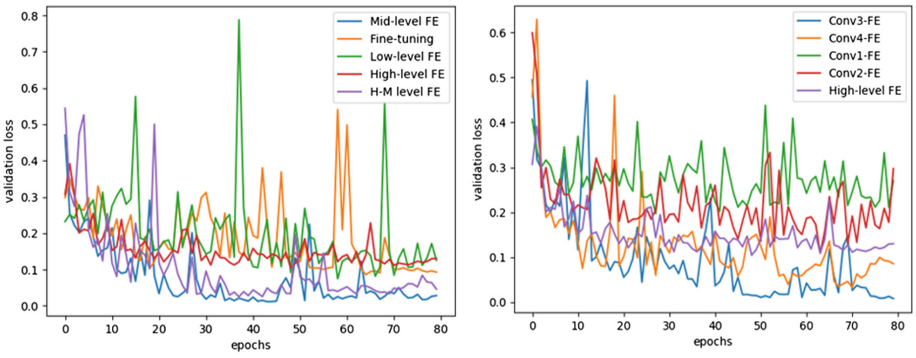


Fig. 2. Validation loss graph by case.

Figure 2 compares each case's validation losses. The mid-level FE shows a stable convergence at 80 epochs; as a result, all the graphs are compared to epoch 80. Figure 2(a) depicts a comparison of each level and fine-tuning and (b) depicts a comparison of each layer. The loss was the lowest and most stable for each mid-level FE comparison to other methods. The mid-level FE shows very good loss performance, demonstrating a maximum of 0.4 and a minimum of 0.02. Furthermore, it displays a stable tendency to converge after 60 epochs. The fine-tuning demonstrates that its learning is unstable and that its loss is high. Because the dataset was small-scale, it

lacked an epoch. We increased the epoch to 200, but they did not converge. The H-M level FE trained in a stable manner, but its loss was higher than that of the mid-level FE. Also, the performance of the high-level FE was lower than each mid-level FE and each low-level FE was poorly trained all the cases, its loss was too high, and it could not be considered to be stable. That is, small-scale medical image datasets are insufficient in the training of low-level layers. Therefore, we confirmed that learning mid-level layers is more effective than learning other levels. This shows that the mid-level FE has the smallest number of epochs to converge, in other words, has lowest computing costs, has the best loss performance, and has a stable learning tendency.

Figure 3 presents comparative ROC curves performed on Montgomery test set. The area under the ROC curve (AUC) of the mid-level FE is 0.87, and the AUC of the fine-tuning is 0.77; in other words, the difference is 0.1, and the accuracy of mid-level FE is higher than the fine-tuning. This demonstrates that our method outperforms the fine-tuning method. Also, it can be observed that the mid-level FE does not overfit the original training dataset.

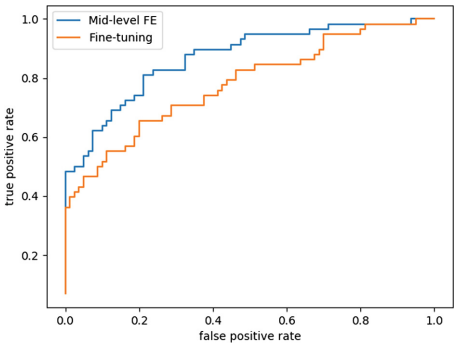


Fig. 3. ROCS of mid-FE and fine-tuning on test set.

We found the mid-level feature extractor to be the most effective method in implementing transfer learning for small-scale medical images. The mid-level feature extractor trained more stably at the mid-level than the model learned at the other levels. In addition, it demonstrated the lowest loss performance and fastest convergence in comparison with the other cases. Therefore, the mid-level feature extractor can be an effective means of training small-scale medical images.

4 Conclusion

We propose a mid-level feature extractor method for the training of small-scale medical imaging datasets. To evaluate the performance of this method, we compared it with low-level FE, high-level FE, and fine-tuning methods. Comparing with other methods, our proposed method shows the lowest amount of loss between 0.4 and 0.02, the most stable training tendency, and the lowest computing costs to converge. In the experiment

pertaining to overfitting, we used different datasets from the training set; the AUC obtained from the test is 0.87. Our method also prevents overfitting; our results are 0.1 higher than the fine-tuning method. Thus, Our method is proved as an efficient alternative to the classification of small-scale medical imaging datasets through its prevention of overfitting and by maintaining its accuracy and reducing computing costs.

Acknowledgments. This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No. 2017-0-018715, Development of AR-based Surgery Toolkit and Applications).

References

1. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998)
2. Bar, Y., Diamant, I., Greenspan, H., Wolf, L.: Chest pathology detection using deep learning with non-medical training. In: Proc. IEEE, 12th International Symposium on Biomedical Imaging, pp. 294–297 (2015)
3. Shin, H., et al.: Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans. Med. Imag. **35**(5), 1285–1298 (2016)
4. Rajpurkar, P., et al.: CheXNet: radiologist level pneumonia detection on chest x-rays with deep learning, arXiv preprint arXiv:1711.05225 (2017)
5. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification with deep convolutional neural networks. In: Proc. NIPS, pp. 1097–1105 (2012)
6. Szegedy, C., et al.: Going deeper with convolutions. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9 (2015)
7. Oquab, M., Bottou, L., Laptev, I., Josef, S.: Learning and transferring mid-level image representations using convolutional neural networks. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 1717–1724 (2015)
8. Razavian, A., Azizpour, H., Sullivan, J., Carlsson, S.: CNN Features off-the-shelf: an astounding baseline for recognition, CoRR, abs/1403.6382 (2014)
9. Zeiler, M., Fergus, R.: Visualizing and understanding convolutional networks. In: Proc. European Conference on Computer Vision, pp. 818–833 (2014)
10. U.S. National Library of Medicine. https://ceb.nlm.nih.gov/repositories/tuberculosis-chest-x-ray-image-data-sets/

Design of an Improved Algorithm for VR-Based Image Processing

Young-Hwan Jang[1], Seung-Su Yang[1], Min-Hyung Park[1], Seungho Han[2], Seok-Cheon Park[3], and Hyungjoon Kim[2(✉)]

[1] Department of IT Convergence Engineering, Gachon University, Seongnam, South Korea
{jang0h, davichi0803, pbh0003}@naver.com
[2] College of Economics and Business Administration, Hanbat University, Daejeon, South Korea
{hansean, hjkim}@hanbat.ac.kr
[3] Department of Computer Engineering, Gachon University, Seongnam, South Korea
scpark@gachon.ac.kr

Abstract. The recent commercialization of virtual reality (VR) has led to various studies in multiple fields that can apply VR technologies, including gaming, design, and virtual experiences. However, when a VR system transmits images through a display located close to the eyes of the user, image output problems, such as image quality degradation and processing speed limitation, could result in a gap between reality and virtual reality, as well as reduce user immersion. Thus, in this study, we have developed an improved image processing algorithm that enhances image sharpness and processing speed to alleviate the user immersion problem in VR environments.

Keywords: Virtual reality · Image processing · Retinex algorithm · Otsu algorithm · Recommendation algorithm · VR image

1 Introduction

The recent commercialization of virtual reality (VR) has led to various studies in multiple fields that can apply VR technologies, including gaming, design, and virtual experiences. These applications of VR technology in various fields will create new economic and social added values beyond ordinary life both domestically and overseas [1, 2]. However, when a VR system transmits images through a display located close to the eyes of the user, image output problems, such as image quality degradation and processing speed limitation, could result in a gap between reality and virtual reality, as well as reduce user immersion. Thus, in this study, we have designed an improved image processing algorithm that reduces the sharpness and size of the output images to resolve the image quality problems in VR environments.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 14–18, 2020.
https://doi.org/10.1007/978-981-13-9341-9_3

2 Related Studies

2.1 Retinex Algorithm

The basic principle of the Retinex algorithm is to remove the background component from the input image. A Gaussian filter is used to extract the background image while ignoring components in the input image whose scales are smaller than the filter size. The Retinex algorithm can be classified into single-scale Retinex (SSR) and multi-scale Retinex (MSR), depending on its application and use. SSR is an algorithm in which a Gaussian filter is used as a single-neighborhood function for a single input image. The equation for the SSR algorithm is as follows.

$$R_i(x,y) = \log\, I_i(x,y) - \log[F(x,y) * I_i(x,y)] \tag{1}$$

$$R_i(x,y) = \log \frac{I_i(x,y)}{F(x,y) * I_i(x,y)} = \log \frac{I_i(x,y)}{\bar{I}_i(x,y)} \tag{2}$$

$$\iint F(x,y)dxdy = 1 \tag{3}$$

In the above equation, I refers to the input image, and F refers to the Gaussian filter. The result obtained from equation is normalized to a value in the range of 0 to 255. The MSR algorithm enhances image quality via multiple neighborhood functions and the weighted sum of these neighborhood functions [3]. The equation for the MSR algorithm is as follows.

$$R_{MSR_i}(x,y) \sum_{n=1}^{N} \omega_n \log I_i(x,y) - \log[F_n(x,y) * I_n(x,y)] \tag{4}$$

2.2 Otsu Algorithm

Image processing speed is determined by the image elements outputted during a specific time. For this reason, unnecessary image elements should be removed to increase image processing speed. The image binarization method can be used for extracting unnecessary image elements. Among the various binarization methods, the Otsu algorithm has been used in various fields owing to its high processing speed and ease of implementation. The Otsu algorithm is a method for determining the threshold that either minimizes or maximizes the variance within classes defined by the sum of the weights of two classes [4]. The equation for the Otsu algorithm is as follows.

$$\sigma_w^2(\mathrm{t}) = \omega_0(2)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t) \tag{5}$$

$$\sigma_b^2(\mathrm{t}) = \sigma^2 - \sigma_\omega^2(t)$$

$$\begin{aligned} &= \omega_0(\mu_0 - \mu_\tau)^2 + \omega_1(\mu_1 - \mu_\tau)^2 \\ &= \omega_0(t)w_1(t)[\mu_0(t) - \mu_1(t)]^2 \end{aligned} \tag{6}$$

In the above equation, the values of weights ω_0 and ω_1 are represented as the probability between the two divided classes, which corresponds to the difference between the two classes σ_0^2 and σ_1^2. This Eq. (6) shows that the inter-class variance maximization is equivalent to its minimization. The image data size and processing speed can be improved by removing the extracted image elements.

3 Improved Algorithm for VR-Based Image Processing

3.1 Overview of the Proposed Algorithm

The most suitable image processing speed for enjoying current VR applications is 16 ms, and the time difference felt by the user is significantly large when the processing speed becomes lower than this time. Furthermore, when the sharpness of the image outputted by the VR device is reduced, the entire image becomes darker. Thus, to solve this problem, this study presents a VR environment suitable for users by removing unnecessary data and improving image sharpness via the proposed algorithm after performing rendering prior to outputting the VR image. Figure 1 shows a schematic diagram of the proposed algorithm.

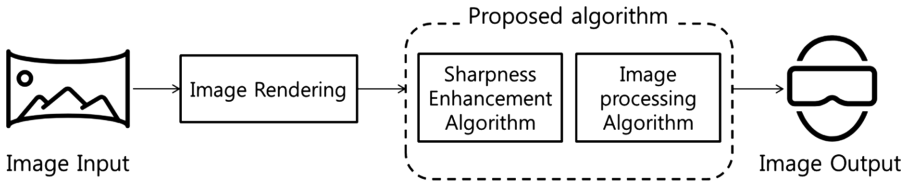


Fig. 1. Overview of the proposed algorithm

3.2 Design of the Proposed Algorithm

In conventional sharpness enhancement algorithms, the sharpness of the input images and photographs are extracted, enhanced, and returned. On the other hand, in the proposed algorithm, light source data are extracted by obtaining a 3D image. These data are modified according to the set weights and are then returned after combining the data with the source images.

The weights applied for enhancing sharpness are obtained using a weighting function, which ensures that the resulting sharpness of the dark regions and the sharpness of the bright regions are balanced via analyses of the detected illumination components. Figure 2 shows the proposed algorithm. The weighting function can suppress the halo effect. The weighting function defines the source image weight as 5,

determines the boundaries of each image, and calculates the divided image boundaries. Furthermore, the weighting function calculates the appropriate weights and applies them to each boundary.

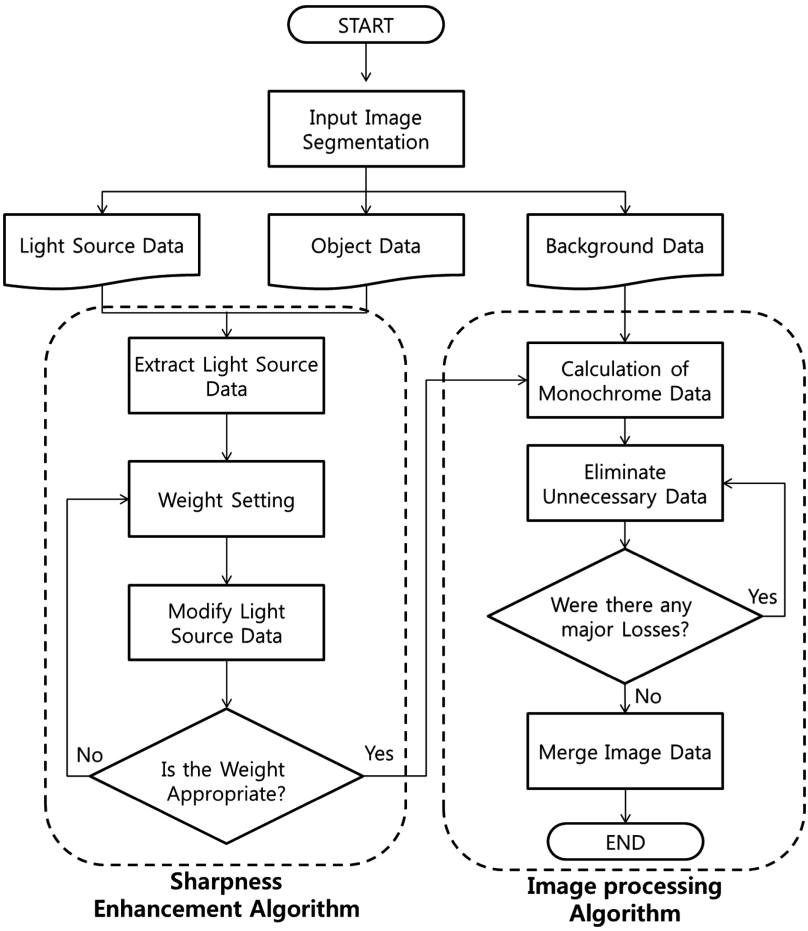


Fig. 2. Proposed Algorithm

VR images are composed of a vast amount of data. Thus, the algorithm operation time becomes longer and the allocated portion of the CPU increases if conventional image processing algorithms are applied, which is problematic. Thus, the proposed algorithm receives 3D images and removes unnecessary data by detecting and comparing grayscale data.

The grayscale data have a size of 640 × 480 pixels and are in the range of 5 to 10 frames per second. Thus, image loss is not significant when removing the grayscale data. Furthermore, the grayscale data are detected without degrading the algorithm's processing speed and with insignificant data loss. The algorithm proposed in this study

simultaneously enhances sharpness and processing speed, which are essential for VR image output. In the final step, the modified image is rendered and displayed to the user.

4 Conclusion

As VR technology emerges as a next-generation technology, various content and technologies based on VR are being developed, further attracting attention to the field.

However, in VR environments, the frame rates should be higher than those of conventional PC and mobile device environments, and the sharpness of the reproduced images should be further enhanced. Thus, in this study, an improved image processing algorithm that reduces the image sharpness and size to resolve image quality problems in VR environments was designed. Future studies will be required to investigate the user recognition of VR images via content creation using the proposed algorithm.

References

1. Song, S.K.: Regulatory reform solution of VR contents industry based on simulator. J. Korea Inst. Inf. Commun. Eng. **21**, 2083–2088 (2017)
2. Jung, J.Y., Sho, K.S., Choi, J.H., Choi, J.H.: Causes of cyber sickness of VR contents: an experimental study on the viewpoint and movement. J. Korea Contents Assoc. **17**, 200–208 (2017)
3. Kang, H.S., Ko, Y.H.: Image quality enhancement method using retinex in HSV color space and saturation correction. J. Korea Multimed. Soc. **20**, 1481–1490 (2017)
4. Lee, Y.W., Kim, J.H.: A computational improvement of Otsu's algorithm by estimating approximate threshold. J. Korea Multimed. Soc. **20**, 163–169 (2017)

Design of Personalized Smart Wellness Information Management System Through Wellness Data Analysis

Young-Hwan Jang[1], Seung-Su Yang[1], Min-Hyung Park[1], Seungho Han[2], Seok-Cheon Park[3], and Hyungjoon Kim[2(✉)]

[1] Department of IT Convergence Engineering, Gachon University, Seongnam, South Korea
{jang0h, davichi0803, pbh0003}@naver.com
[2] College of Economics and Business Administration, Hanbat University, Daejeon, South Korea
{hansean, hjkim}@hanbat.ac.kr
[3] Department of Computer Engineering, Gachon University, Seongnam, South Korea
scpark@gachon.ac.kr

Abstract. Recently, the health concept is changing from simply being healthy to a proactive perspective of balancing physical and psychological factors. Consequently, interest in health-oriented wellness is increasing. Wellness, which can improve quality of life through continuous recording and management, is also continuously advancing. However, at present, classification and implementation of major wellness information in daily life has not yet been achieved. Therefore, in this study, a personalized smart wellness information management system is designed.

Keywords: Wellness · Management system · Smart wellness · Data analysis · Recommendation algorithm · Wellness contents

1 Introduction

Recently, the health concept has begun to socially change from health simply being about not having diseases to a proactive perspective of balancing physical, psychological, and intellectual factors [1]. Accordingly, interest in health-oriented wellness is increasing. The term wellness refers to the combination of well-being, happiness, and fitness; it also refers to a physically, psychologically, and socially healthy state [2]. Furthermore, at the World Health Organization, the health and wellness concepts are changing from simply being the opposite of the illness to a more proactive perspective.

As the wellness perspective is changing, its importance in daily life is also increasing. Moreover, by quantifying information related to daily activities, people can improve their quality of life through continuous recording and management [3]. Wellness data are appropriate information derived by recording and analyzing all the data related to an individual's life, and through integration with the latest information and communications technology (ICT), it can change personal life positively by

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 19–24, 2020.
https://doi.org/10.1007/978-981-13-9341-9_4

supporting a smart life environment construction. However, classifying and implementing major wellness information in daily life has not yet been achieved. Therefore, a smart wellness information management system is proposed here to systematically analyze and monitor a user's wellness records and provide customized service.

2 Related Studies

2.1 Wellness

Traditionally, health was perceived as a state of no disease or no physical and psychological weaknesses. However, recently, the health concept has included wellness, thus encompassing all physical and psychological factors. Wellness is a new concept that helps people change their lifestyle for optimum health and the wellness domain is classified into physical, psychological, and emotional subdomains [4].

Wellness was first mentioned in 1654 as an antonym of illness or disease, and now it means a multi-dimensional approach to healthy life, as a compound word indicating wellbeing and fitness. Wellness content refers to content used to support a user's multi-dimensional healthy life [5]. Particularly, the ICT-fused medical industry has emerged as a core industry in the 4th industrial revolution era, and it encompasses services and systems that provide personal health condition management and customized medical and health management regardless of time and location. According to the classification standard, health includes smart healthcare, U-health, wellness, and more [6].

3 Design of Personalized Smart Wellness Information Management System

3.1 Overview of the Proposed System

The system proposed in this paper stores all information related to the user's daily life and provides information search and management functions. The proposed system extracts only the wellness information designated as standard data through the designed information management algorithm based on the information inputted by user. The extracted information is used to provide daily life pattern analysis and information recommendation to the user through the information management algorithm and the user can manage their wellness information systematically by receiving appropriate daily life guidance. An overview of the proposed system is illustrated in Fig. 1.

3.2 Proposed System's Structural Diagram

The personalized smart wellness information management system proposed in this paper uses personalized records, which have been reclassified according to the wellness value; it then provides them to the user through information extraction. The proposed system's overall structure diagram is shown in Fig. 2. The data inputted through a smartphone by user are transmitted to the proposed system. The transmitted data are used to extract wellness information through the wellness information management

algorithm, and they are stored in the database. Furthermore, when the user requests, the user's analyzed wellness information is saved in the database; the generated wellness information is delivered to the user through their smartphone.

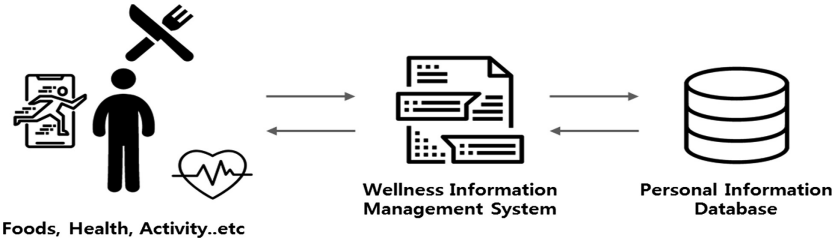


Fig. 1. Overview of the proposed system

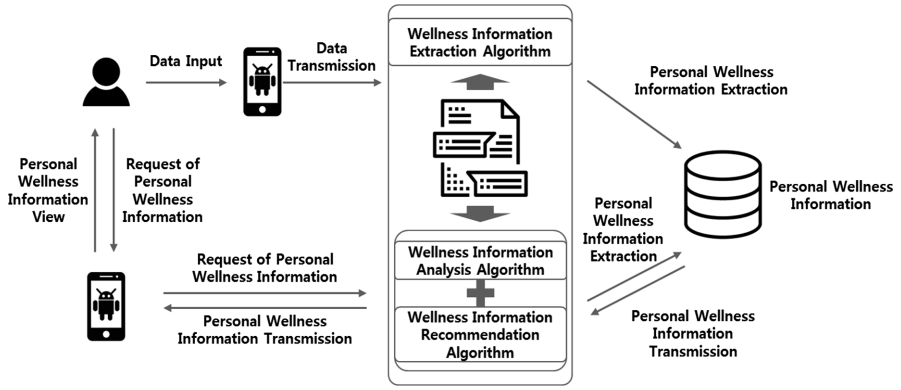


Fig. 2. Proposed system's overall structure diagram

3.3 Design of Wellness Information Management Algorithm

The proposed personalized smart wellness information management system is designed in such a way that the data input by a user through a smartphone can be converted and extracted as important wellness information via the information management algorithm. Moreover, it was designed to facilitate a wellness service to the user by providing daily life pattern analysis and personalized daily life guidance based on the user's wellness information. The algorithm is shown in Fig. 3.

The designed algorithm has three stages: wellness information extraction, analysis, and recommendation. In the wellness information extraction stage, it is checked whether the input data is standardized, and if it is, it is saved. After checking all input data, the final wellness information is acquired and the process is terminated. The wellness information extraction algorithm is shown in Fig. 4.

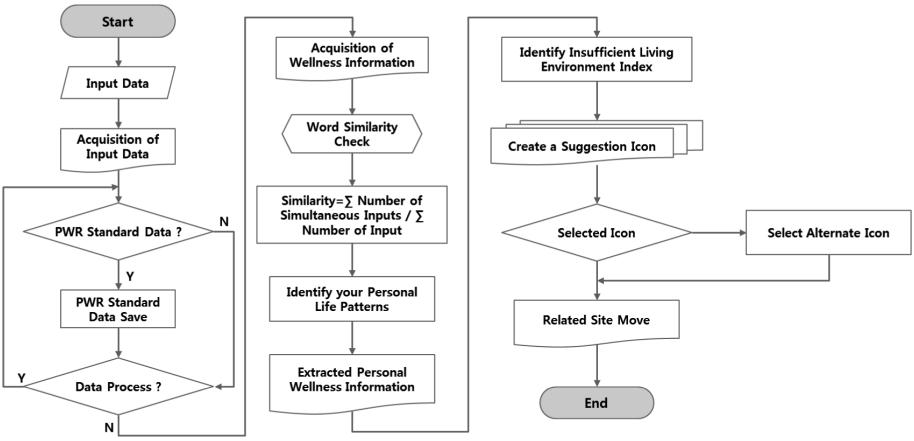


Fig. 3. Overview of the full algorithm

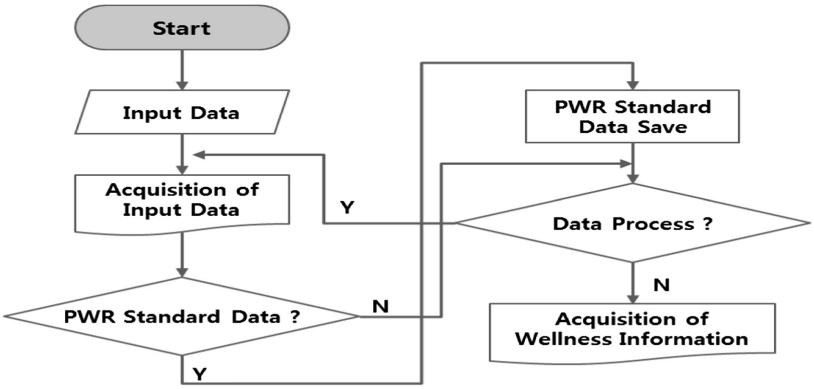


Fig. 4. Wellness information extraction algorithm

After extracting the wellness information, it is analyzed, and in the system, the similarity degree between words is calculated through Eq. (1). In Eq. χ and γ refer to the words, and the similarity degree between the user's wellness information is verified through Eq. (1). Using the verified result, the user's daily activity analysis and wellness information are extracted. The algorithm that analyzes the wellness information is shown in Fig. 5.

$$W_{xy} = \frac{\sum(Number\ of\ times\ x\ and\ y\ were\ entered\ together)}{\sum(Number\ of\ times\ x\ or\ y\ was\ entered)}$$

The wellness information recommendation algorithm was designed to provide and recommend appropriate daily life guidance to users based on the extracted wellness information. Through the recommended features option, it is possible to go to a related

website and obtain a different recommendation based on user choice. The recommendation algorithm is shown in Fig. 6.

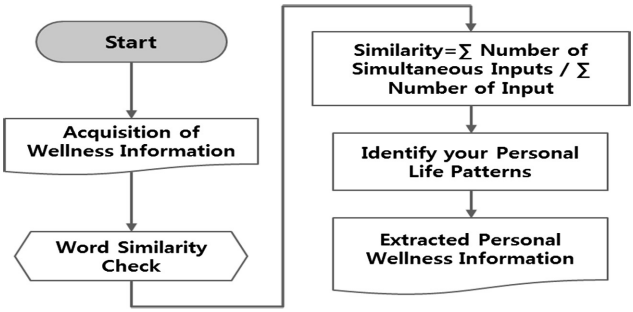


Fig. 5. Wellness information analysis algorithm

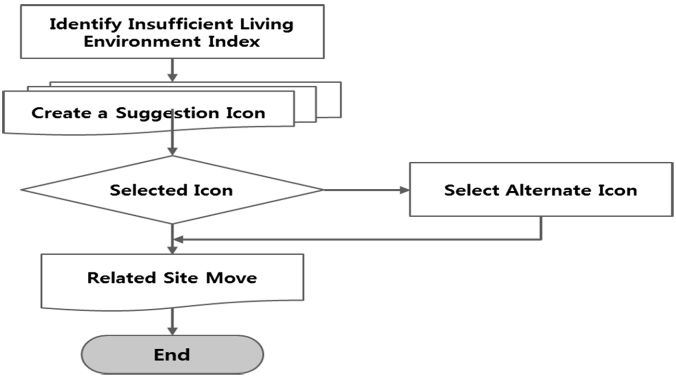


Fig. 6. Wellness recommendation algorithm

4 Conclusion

Recently, as interest in health is increasing, the interest in wellness to pursue a better quality of life is increasing rapidly. Furthermore, the past passive consumer behavior of receiving simple services is transitioning to active consumer behavior using smartphones and similar devices. According to such changes, the health management service and technology provided through smartphones are expected to advance in various health-related areas including wellness.

In this study, a personalized smartphone management system was designed and implemented through a wellness information management algorithm that uses user's wellness records. Prior to designing the system proposed in this paper, the system requirements were identified and defined, and the wellness information management algorithm was designed to extract wellness data according to the defined standard.

Furthermore, the wellness information recommendation algorithm was designed to provide appropriate daily life guidance or patterns to the user.

References

1. Ahn, H.S.: Relationships between wellness, campus-life satisfaction and learning flow among university students. J. Korea Contents Assoc. **16**, 494–502 (2016)
2. Lee, W.J.: A study on recommendation service system for the customized convergence wellness contents. J. Korea Multimed. Soc. **20**, 322–329 (2017)
3. Lee, W.J.: A study on the design and implementation of service system for the convergence wellness content in mobile environments. J. Korean Inst. Inf. Technol. **14**, 155–163 (2016)
4. Jeon, B.K., Ahn, H.C.: Collaborative filtering app recommendation system using user review mining. P. Korea Intell. Inf. Syst. Soc. 25–44 (2015)
5. Lee, W.J.: Designing of the dynamic user profile for wellness content recommendation based on user's feedback. J. Korean Inst. Inf. Technol. **14**, 149–158 (2016)
6. Choi, M.J., Kim, S.H.: Relationship between wellness components and wellness status. P. Korea Intell. Inf. Syst. Soc. 229–243 (2015)

Design of POI Extraction Speed Improving Algorithm Based on Big Data

Young-Hwan Jang[1], Seung-Su Yang[1], Min-Hyung Park[1], Seok-Cheon Park[2], and Hyungjoon Kim[3(✉)]

[1] Department of IT Convergence Engineering, Gachon University, Seongnam, South Korea
{jang0h, davichi0803, pbh0003}@naver.com
[2] Department of Computer Engineering, Gachon University, Seongnam, South Korea
scpark@gachon.ac.kr
[3] College of Economics and Business Administration, Hanbat University, Daejeon, South Korea
hjkim@hanbat.ac.kr

Abstract. Location-based services (LBSs) that collect and utilize location data in real time through mobile devices have been widely used recently. LBSs use a clustering algorithm to extract a user's point of interest (POI) from a stay point; the POI is defined as a place that an individual stays in or uses for a given amount of time. However, the DBSCAN algorithm increases the amount of unnecessary iterative clustering computations as the amount of stay point data increases, thus causing an increase in overall computation time. Therefore, in this paper, we propose an algorithm to improve big data-based POI extraction speed.

Keywords: Big data · POI · DBSCAN · Stay point · Clustering algorithm · LBS

1 Introduction

The recent widespread use of various mobile devices, such as smart phones and tablets, has increased the volume of location data that is generated in real time. As a result, location data can be utilized in various fields such as user location detection and movement pattern analysis, and the number of services based on location data is increasing. In addition, in order to utilize the available location information, a point of interest (POI) where many users often stay should be extracted from a stay point (sp) containing location information [1, 2].

Various clustering techniques are used to extract POI information. The DBSCAN algorithm is widely used among such density-based clustering methods. However, algorithms for density-based clustering methods have the same number of iterations as the number of data, which means the time complexity increases as the number of stay point data increases. Therefore, in this paper, we propose an algorithm to reduce the

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 25–29, 2020.
https://doi.org/10.1007/978-981-13-9341-9_5

number of computations needed for large volumes of location data and efficiently extract POI information by improving the DBSCAN algorithm structure.

2 Related Studies

2.1 Location-Based Service (LBS)

Location-based services (LBSs) measure location using only GPS, GNSS, and wireless mobile communications in terminals while moving. They also provide various application services required by users based on their location. In particular, with the increase in use of mobile devices with built-in GPS modules, such as smart phones and tablets, there is a growing interest in LBSs that provide services by collecting user or device location data. LBSs are mainly classified as follows: positioning technology, location processing platform, or location processing program [3].

2.2 Point of Interest (POI)

A POI is a place that people often stay in or use. It refers to a specific location on a map that can be useful or interesting to people. A place that many users continuously visit may be a meaningful place and contain meaningful information; thus, the information about such a place becomes essential for the location information services.

POIs can be displayed in an electronic map service, such as satellite navigation or Google Maps and are mainly divided into two categories based on their attributes.

First, there are POIs containing essential information such as area name, address, contact information, coordinates, and building information and containing detailed information such as a parking lot status or photographs of a certain location. These are POIs with various attributes but without location information. If we select a specific location when searching a map, detailed information about the location is available to easily identify its suitability. Second, there are POIs displaying only basic location information such as latitude, longitude, and time. They contain only the coordinates and time measured by satellites and can be used to track location movement in real time or to check movement trajectory over time [4].

3 Design of POI Extraction Speed Improving Algorithm

3.1 Overview of the Proposed Algorithm

The proposed algorithm selects the initial cluster in the same way as the existing DBSCAN method and then selects an sp to execute the cluster extension algorithm. If we search for an sp with the longest distance without any condition, then it may be impossible to expand in a certain direction depending on the sp distribution in the cluster. Therefore, the cluster range within the distance threshold is divided into quadrants based on the core sp. After this, the distance between the core sp and the neighboring sp in the cluster is calculated in each quadrant using the Euclidean distance metric. The equation for the Euclidean distance is shown in Fig. 1.

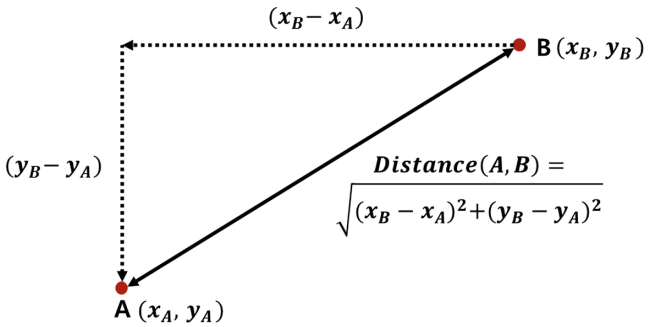


Fig. 1. Equation for the Euclidean distance

By checking all sps in the cluster using the distance measurement method, the sp included within the distance threshold becomes the neighboring sp. In each quadrant, the sp with the longest distance from the core sp is the far-sp (f-sp) for executing the algorithm when expanding the cluster. The sp selected for executing the algorithm is shown in Fig. 2.

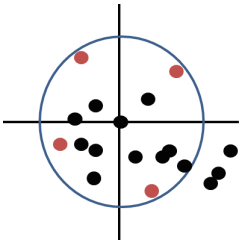


Fig. 2. SP selected for executing the algorithm

However, if the clustering algorithm is executed only for the f-sp as shown in Fig. 2, the proposed algorithm can miss the sp in the cluster because it is not included in the distance threshold in which the algorithm operation is executed. To minimize this problem, the distance is measured by connecting the selected f-sps in neighboring quadrants (Q1–Q2, Q2–Q3, Q3–Q4, and Q4–Q1) and deriving the median value of the distance.

After that, the cluster blank is minimized by finding and selecting the sp that is closest to the median value and farthest from the core sp, thus improving the accuracy by reducing the missing sp. Figure 3 shows how to find the sp that is farthest from the core sp based on the median value of the f-sp. Through this process, sp data can be analyzed and converted into POI data. Figure 4 shows the overall process diagram of the proposed algorithm.

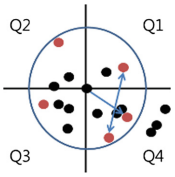


Fig. 3. Core SP based on the median value of the f-sp

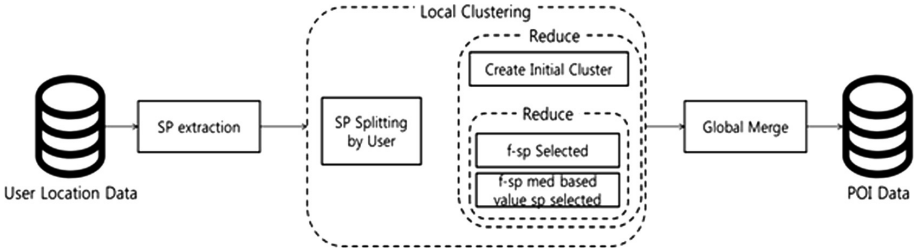


Fig. 4. Overall process diagram of the proposed algorithm

3.2 Design of the Proposed Algorithm

There is a shortcoming with existing clustering algorithms: the larger the data density, the longer is the overall algorithm execution time. The clustering algorithm is not suitable owing to the nature of big data analysis systems, which must analyze and process large amounts of data quickly. The proposed clustering algorithm improves the overall algorithm execution time by decreasing the number of operations for each point in the DBSCAN algorithm. Figure 5 shows the proposed algorithm.

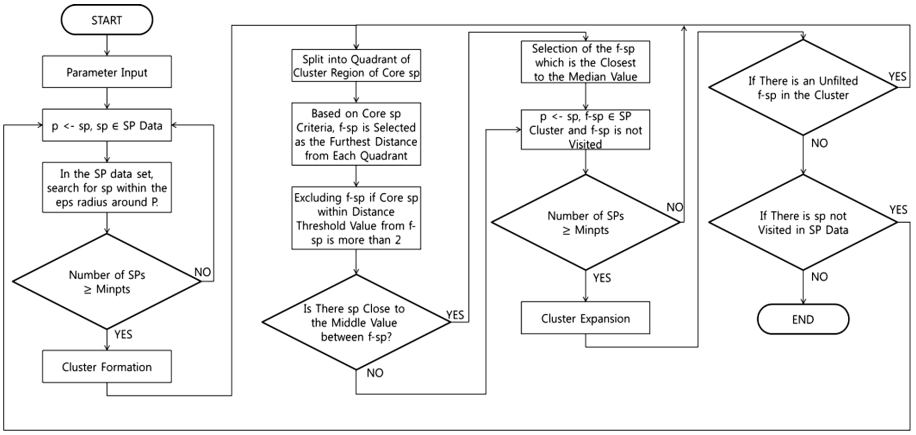


Fig. 5. Overall process diagram of the proposed algorithm

After forming the initial cluster, it divides the cluster into quadrants and finds the distance between core sp and sp in each quadrant. After this, it improves the clustering speed by selecting the minimum number of f-sps for performing a cluster expansion operation. Lastly, the output POI data are saved into the database.

4 Conclusion

With the widespread use of GPSs due to recent advancements in mobile technologies, studies have been conducted to analyze movement patterns using user movement trajectory and identifying user POIs based on analysis; this can then be applied to LBS.

We proposed a method to extract POIs using the density-based clustering algorithm DBSCAN, which enables clustering arbitrary shapes and processing noise data better than various other clustering algorithms used in previous research. However, as the DBSCAN algorithm performs a clustering operation for all sps in the cluster, the computation time is large owing to the unnecessary computation of duplicate sps. Therefore, in this paper, we proposed an algorithm to improve POI extraction speed by improving the clustering algorithm structure.

References

1. Lee, J.E., Son, H.M., Yang, J.H., Yu, K.Y.: Design and implementation of semantic search for POI utilizing collective intelligence. J. Korean Soc. Surv. **34**, 339–346 (2016)
2. Shin, W.Y., Vu, D.D.: Density-based estimation of POI boundaries using geo-tagged tweets. J. Korean Inst. Commun. Inf. Sci. **42**, 453–459 (2017)
3. Jiang, T.P., Lim, H.A., Choi, J.W.: The effectiveness of apps recommending best restaurant through location-based knowledge information: privacy calculus perspective. J. Soc. e-Bus. Stud. **22**, 89–106 (2017)
4. Shin, W.Y., Choi, S.I.: Improved estimation of social POI boundaries through joint optimization. J. Korean Inst. Commun. Inf. Sci. 2046–2049 (2017)

Analyzing Twitter Data of Family Caregivers of Alzheimer's Disease Patients Based on the Depression Ontology

Hyon Hee Kim, Sohee Jeong, Annie Kim, and Donghee Shin(✉)

Dongduk Women's University, 60 Hwarang-ro, 13-gil, Seongbuk-Gu, Seoul 136-714, South Korea

Abstract. In this paper, we present an ontology-based approach to analyze depression of family caregivers of Alzheimer's disease patients. First, we developed depression ontology called OntoDepression considering the language written in social media. Four major classes and specialized subclasses are defined based on the dailyStrength, which is a well-known social media site centered on healthcare. Next, to find mental health of family caregivers of Alzheimer's patients, their twitter data is analyzed based on the OntoDepression. Our experimental results show that negative feelings of family caregivers are not clearly revealed, while medical condition of depression symptom is highly rated. Also, their tweets mention a lot about human relationships, work and activities.

Keywords: Tweet analysis · Depression ontology · Family caregivers · Alzheimer's disease · Mental health of family caregivers of AD patients

1 Introduction

As we enter the aging society, the prevalence of Alzheimer's disease (AD), a geriatric disease, is rapidly increasing. In particular, since the number of patients with Alzheimer's disease increases by two times every five years after age 65, in Korea, it is estimated to reach one million in 2030 and two million in 2050 [1]. Therefore, family caregivers of AD patients are becoming important worldwide. It is well-known that the stress and depression of family caregivers of AD patients, often called the invisible second patients, is severe [2–4]. To identify Alzheimer's caregiver stress, concept analysis is performed [5]. By analyzing the related literature, antecedents and consequences of caregiver stress are defined. Consequences are as follows: depression, anxiety, irascibility, cognitive disturbance, poor health status, yielding role, and role entrapment due to guilt and shame of caregivers.

Early depression studies have been done based on the user study or surveys. To assess depression scales and criteria, depression assessment scale is developed such as Beck's Depression Inventory [6] and Hamilton Depression Rating Scale [7]. These criteria have been applied to real-world cases for many years. More recently, detection and prediction of depression via social media have attracted much attention [8]. In these days, lots of people use social media every day, postings of users include their

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 30–35, 2020.
https://doi.org/10.1007/978-981-13-9341-9_6

everyday life, feeling, mental condition, and so forth. Therefore, social media is a valuable resource to capture depressive moods of users. In particular, analyzing tweet postings in Twitter showed successful results to predict depression [9, 10]. Therefore, research on depression detection via online social media data is essential.

In this paper, a tweet analysis of family caregivers of AD patients is performed to identify their mental health and to understand their life and feelings. For this purpose, Ontology for depression, called OntoDepression, is developed based on postings of DailyStrength [11] which is a social media for support groups. In the website, users share their information about disease and discuss their struggles and successes with each other. From the depression category, the total number of 1012 postings is collected and 1,000 important keywords are extracted using TF-IDF algorithm. We adopt 4 main classes such as symptoms, treatments, feelings, and life defined in [9], and specialized symptom class and life class in detail. The symptom class is classified into medical condition, physical condition, mental condition, and general condition subclasses. The life class is classified into work & activities, human and alive things, time, location, related to human body parts, general words subclasses. The feeling class is simply classified into negative and positive subclasses.

The contribution of this paper is in the development and validation of the ontology-based framework for depression diagnosis using online social media. To collect tweets of family caregivers of AD patients, we extracted twitter data with hashtags "Alzheimer's and mom", "Alzheimer's and dad", "Alzheimer's and wife", and "Alzheimer's and husband". The total number of 9,486 tweets is analyzed. Our experimental results show that unlike our expectations, negative feelings of family caregivers of AD patients did not show significant difference from the positive feelings. On the other hands, it is noteworthy that keywords representing medical condition are highly rated. Also, family caregivers of AD patients mainly mention human & alive things and work & activities.

The remainder of this paper is organized as follows. In Sect. 2, we describe development of depression ontology. In Sect. 3, we explain the experimental results of tweets analysis using the framework. Finally, in Sect. 4, we give concluding remarks.

2 Development of Depression Ontology

Figure 1 shows the overview of the ontology-based framework for depression analysis. Basically, we defined four main classes considering related literature and extended the main classes with the specialized features. For this purpose, important keywords from social media, dailyStrength, are extracted and are used to define subclasses. To analyze depression of family caregivers of AD patients, tweets are collected. Retweets, tweets for advertising and tweets from official sites related to Alzheimer's such as Alzheimer's society and Alzheimer's Association are also removed. In addition, to extract tweets from family caregivers with AD patients, keywords by combining Alzheimer's and mom, dad, wife, or husband are used. By analyzing the twitter data based on OntoDepression, the state of depression is explained.

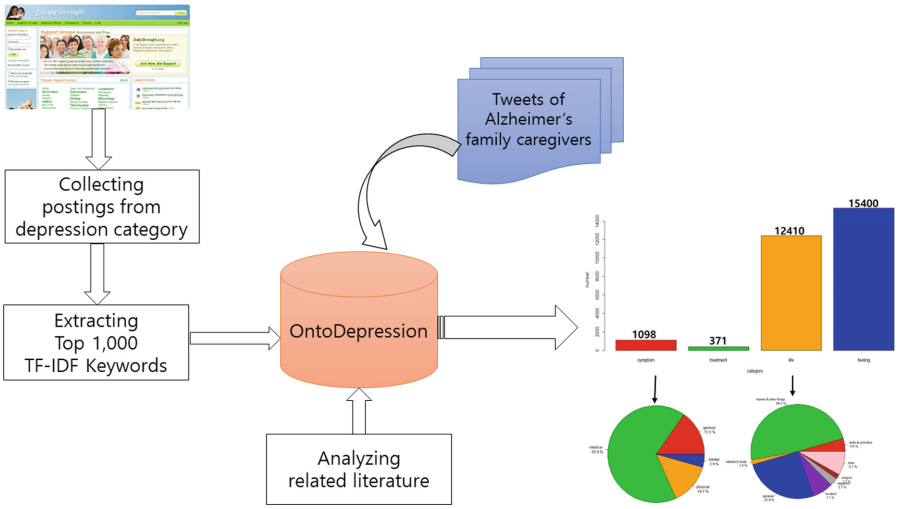


Fig. 1. Overview of the framework

Table 1 shows the structure of OntoDepression and examples of instances in detail. OntoDepression has four classes, *symptoms*, *treatments*, *feelings*, and *life*. The class *symptoms* indicates the specific symptoms of depression, and has four specialized subclasses, *medical condition*, *physical condition*, *mental condition*, and *general condition*. The class *treatments* represents concept of medical treatments such as drugs,

Table 1. Structure and instances of OntoDepression

Class	Subclasses	Examples of instances	%
Symptoms	general	additive, imbalance, alcoholism, chronic	35%
	medical	bipolar, dementia, asthma, thyroid, agoraphobia	13%
	physical	awake, dizziness, fever, headache, insomnia	14%
	mental	delusion, fantasize, flashbacks, ghost, monster	38%
Treatments	–	antidepressant, clinic, drugs, medication, pills	100%
Feelings	negative	anxious, confused, abusing, cheated, destroyed	70.3%
	positive	acceptable, thankful, interesting, encouraging	29.7%
Life	human & alive things	adult, aunt, baby, brothers, boyfriends, cat, dog, camel, cousin, coworkers, father, fiancé, fish	12%
	work & activities	dating, cooking, event, game, gardening computer, companies, homework, hobbies	8%
	body parts	ankle, arm, ass, brain, eyes, skin, teeth, throat	2.4%
	location	army, hell, high school, home, grave, forest	4.4%
	time	afternoon, birthday, holidays, midnight, Easter	4.4%
	negative things	accident, assaulted, blocked, harming, injury	9%
	religion	bible, Jesus, Christ, lord, pray, grace, god, prayer	2%
	general	email, everything, dishes, dreams, flight, gift	58%

medication, and pills. The class *feelings* is classified into two subclasses, *negative feelings* and *positive feelings*. Finally, the class *life* captures what the family caregivers are saying about. As shown in Table 1, it has 8 subclasses such as *human & alive things*, *work & activities*, *human body parts*, *location*, *time*, *negative things*, *religion*, *and general things*. The subclasses of the class *life* are defined based on the questions of Beck's depression inventory [6] and Hamilton Depression Rating Scale [7]. The subclass *human & alive things* captures social relationship of family caregivers. The subclass *work & activities* describes what they are doing, and the subclass *negative things* describes negative words related to everyday life.

Algorithm 1 explains how to build OntoDepression. The postings from depression category are read and converted into a corpus (lines 8–9). The corpus is cleaned by preprocessing (lines 10–15). The keyword list is built based on term frequency-inverse document frequency weight [12] (line 18). The class hierarchy is also built, and the keywords in the keyword list are registered in the specified classes as instances. Executing the function buildDictionary(), OntoDepression is built (line 27). In the case of class *feelings*, terms in an existing sentiment dictionary [13] are added (line 28). The total number of 7,649 terms is processed in OntoDepression. OntoDepression is implemented using R, a statistical computing language in windows environment.

Algorithm 1. Building OntoDepression

```
1: begin initialization
2:      dsRawData              null;
3:      dsCorpus               null;
4:      dsCorpusClean          null;
5:      OntoDepression         null;
6: end
7: begin preprocessing
8:      dsRawData <- Read Data;
9:      dsCorpus <- Convert dsData into dsCorpus;
10:     while not untill end of line of the dsCorpus do
11:           Remove special character;
12:           Remove numbers;
13:           Remove stopwords;
14:           dsCorpusClean <- dsCorpus;
15:     end
16: end
17: function buildDictionary()
18:     keywordList <- Find top n terms with TF-IDF weight;
19:     classList <- Assign the name of classes, Symptoms, Treats, Feelings, and life;
20:     subclassList <- Assign the names of subclasses;
21:     Make superclass-subclass hierarchy;
22:     while not untill end of line of the keywordList do
23:           Register a keyword to a specified subclass;
24:     end
25: end
26: begin create OntoDepression
27:     OntoDepression <- buildDictionary();
28:     OntoDepression <- Add sentimentDictionary as instances of Feelings class;
29: end
```

3 Analysis of Twitter Data

To validate the OntoDepression and to identify depression state of family caregivers of AD patients, tweets analysis is performed based on OntoDepression. From January 1, 2017 to December 31, 2017, the number of 9,611 tweets from family caregivers is collected. To extract family caregivers' tweets among lots of tweets related to Alzheimer's disease, only tweets containing the following words like mom, dad, wife, and husband are chosen. As a result, the number of 9,486 tweets are analyzed. Figure 2 shows the result of depression analysis. First, among the four major classes, the class *feelings* contains terms over 50%, and the class *life* contains 42.2% terms. The class *symptoms* and the class *treatments* are 3.7% and 1.3%, respectively. It reveals that tweets of family caregivers of AD patients mainly mention their daily life. The specific symptoms and treatments of depression account for about 5% of the total.

Unlike our expectation, negative feelings (52.7%) do not appear significantly more than positive feelings (47.3%). This result can be interpreted as two ways. One is that family caregivers of AD patients may not express their deep emotions with a short message containing only 120 characters' limit. The other is that they actually do not have negative feelings. In the case of the class *life*, the subclass *human & alive things* is highly rated as 48.3%. Tweets belonging to the subclass *human & alive things* mainly mention relationships with other people like family, friends, and coworkers or pet animals. Their tweets mention time (8.1%), location (7.1%), and also work & activities (4.6%). In addition, their tweets mention negative aspects of life (4.6%), religion (1.7%), and body parts (1.6%). The class *symptom* contains medical symptoms (65.9%), general symptoms (15.5%), physical symptom (14.7%), and mental symptom (3.9%). It is noteworthy that medical symptoms are highly rated in the class *symptom*.

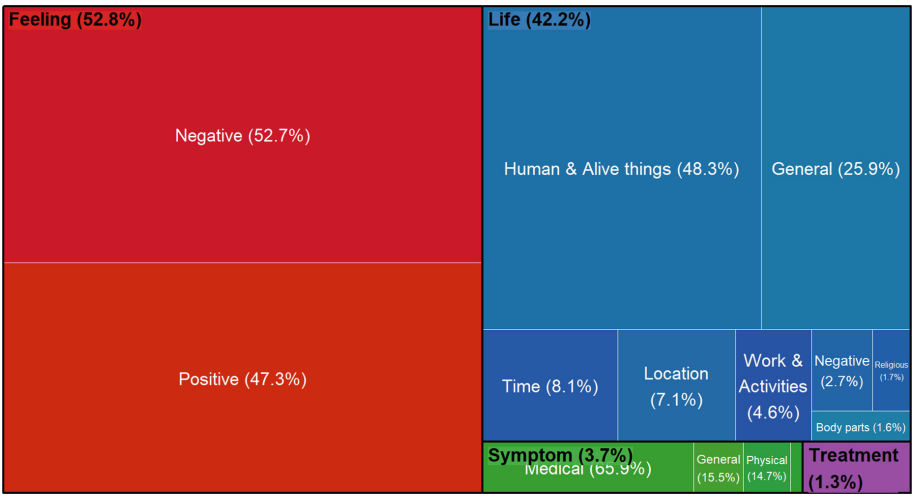


Fig. 2. Result of depression analysis

4 Conclusions

In this paper, we present an ontology-based approach for depression analysis using social media. Because language habits used in social media reflect people's depression, an ontology for depression analysis called OntoDepression is developed based on postings belonging to depression category in a social media site. To test the OntoDepression ontology, family caregivers of AD patients which are rapidly growing are targeted. Tweets of family caregivers of AD patients are collected and analyzed.

Our research results show that family caregivers of AD patients' negative feelings are not directly revealed on their tweets, but their tweets contains medical symptoms of depression a lot. To understand this more specifically, research on collecting and analyzing long-term tweets of individual users is currently underway. They mention human relationships with family, neighbors, friends, and coworkers a lot and also mention general life, time, location, work & activities in order. We are currently extending the OntoDepression by specializing the subclass *negative feeling* and the subclass *positive feeling*.

References

1. Kim, S.H., Han, S.: Prevalence of dementia among the South Korean population. J. Korean Diabetes **13**(3), 124–128 (2012)
2. Lavretsky, H.: Stress and depression in informal family caregivers of patients with Alzheimer's disease. Aging Health **1**(1), 117–133 (2005)
3. Brodaty, H., Donkin, M.: Family caregivers of people with dementia. Dialogues Clin. Neurosci. **11**(2), 217–228 (2009)
4. Mahoney, R., Regan, C., Katona, C., Livingston, G.: Anxiety and depression in family caregivers of people with Alzheimer's disease. Am. J. Geriatr. Psychiatry **13**(9), 795–801 (2005)
5. Lianque, S., et al.: Concept analysis: Alzheimer's caregiver stress. Nurs Forum **51**(1), 21–31 (2015)
6. Beck, A., Mendelson, M., Mock, J., Erbaugh, J.: An inventory for measuring depression. Arch. Gen. Psychiatry **4**(6), 561–571 (1961)
7. Hamilton, M.: A rating scale for depression. J. Neurol. Neurosurg. Psychiatry **23**, 56–62 (1960)
8. Shen, G., et al.: Depression detection via harvesting social media: a multimodal dictionary learning solution. In: Proceedings of the 26th International Joint Conference on Artificial Intelligence, pp. 3838–3844 (2017)
9. Choudhury, M., Gamon. M., Counts, S., Horvitz E.: Predicting depression via social media. In: Proceedings of the 7th International AAAI Conference on Weblogs and SocialMedia, pp. 128–137 (2013)
10. Park, M., Cha, C., Cha, M.: Depressive moods of users portrayed in Twitter. In: Proceedings of the 18th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (2012)
11. DailyStrength. http://www.dailystrength.org/
12. Manning, C.D., Raghavan, P., Schutze, H.: Introduction to Information Retrieval, 1st edn, pp. 116–121. Cambridge University Press, New York (2008)
13. Sentiment Dictionary. https://github.com/jeffreybreen/twitter-sentiment-analysis-tutorial-201107

Research on the Automatic Classification of Ship's Navigational Status

Jaeyong Oh(✉), Hye-Jin Kim, and Sekil Park

Korea Research Institute of Ships and Ocean Engineering, 32 1312 Beon-Gil, Yuseong-Daero, Yuseong-Gu, Daejeon, Republic of Korea
{ojyong, hjk, skpark}@kriso.re.kr

Abstract. Maritime traffic analysis has been attracted increasing attention due to their importance for the safety and efficiency of maritime operations. The first step of maritime traffic analysis is the identification of ships' navigational status, and various analysis tasks are started based on the status information. It should be considered the complex traffic characteristics of the harbor and ships. These tasks depend on the expert's experiences, however, it becomes difficult to classify manually as the amount of traffic volume increases. Therefore, in this paper, we proposed a new model to identify the ship's navigational status automatically. The proposed method generated traffic pattern model using accumulated AIS trajectories and then classified using the clustering algorithm. This method based on semi-supervised machine learning and the proposed clustering method using the pre-classified dataset. Finally, we review experimental results using the actual trajectory data to verify the feasibility of the proposed method.

Keywords: Navigational status · AIS · Machine learning · Clustering

1 Introduction

The identification of the ship's status is the first step for maritime situation awareness, also it is a significant issue of maritime safety. The navigational status information can be used for VTS (Vessel Traffic Services) or maritime traffic analysis tasks as well as navigational safety. The AIS (Automatic Identification System) is a mandatory navigational device. And, the AIS signals include a Navigational Status field which is reported by the vessel. However, in some cases, the status information has delayed updates or navigate without updating because the information manually set by the crew. It makes errors in the results of traffic analysis using wrong status information. Also, it affects safety navigation of ships. Therefore, in this paper, we propose an automatic classification model of ship's navigational status using the AIS trajectory data. We propose the machine learning model for generating traffic patterns and describe how to make a training data. Also, we introduce a recursive clustering method for status classification. Across experiments with AIS trajectory data, we review the feasibility of the proposed method.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 36–41, 2020.
https://doi.org/10.1007/978-981-13-9341-9_7

2 Related Works

The IMO (International Maritime Organization) requires AIS to be fitted aboard international voyaging ships with 300 or more gross tonnage (GT), and all passenger ships regardless of size [1]. AIS messages include various navigational information for example position, course, speed, etc. The main purpose of AIS is collision avoidance, and many other applications have developed. AIS is currently used for VTS, aid to navigation and maritime traffic analysis. There is information of navigational status in the AIS messages, and it consists of 8 different statuses. The navigational status can be set to "underway", "anchoring", "mooring" or "fishing" under the normal situation. And, in the case of emergency, it is available to set "aground" or "not under command". However, there is some information that requires manual entry such as navigational status. In some cases, it has a different status between AIS messages and actual vessel. Figure 1 shows the navigational status from AIS messages. Although there are some trajectories in "underway" status, in fact, it is "mooring" status, and the opposite is the case also.

Numerous studies have been conducted on maritime situation awareness using ship traffic data. Perez et al. introduced a method for estimating vessel emissions from AIS trajectories using spatial analysis method [2]. Gloaguen et al. showed an effective way to identify and to predict fishing activity in AIS trajectories using data mining and machine learning techniques such as HMM(Hidden Markov Model) [3, 4]. Pallotta et al. proposed an unsupervised incremental learning approach to extract traffic route and detect anomalies. This method based on statistical algorithms, also it is difficult to identify the change of the status during navigation because the route is extracted from entry and exit points [5].

In these studies, trajectory patterns are used to identify the type of ship or to detect specific activities such as fishing. However, there is a problem that it is difficult to define classification criteria and to generate automatically a large training set for supervised learning.

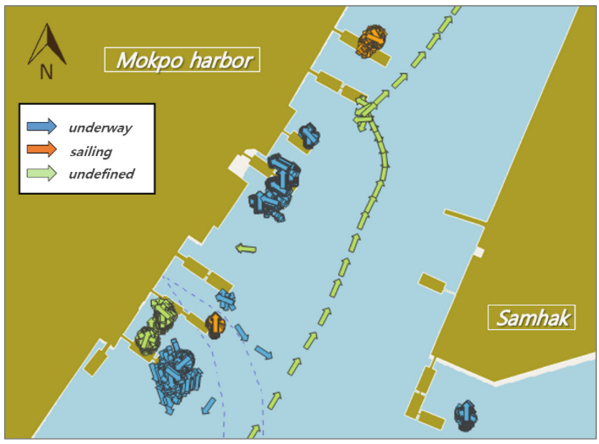


Fig. 1. Navigational status in AIS messages

3 Classification Model

3.1 Concepts

In this paper, we proposed a model that identifies the ship's navigational status automatically by training the pattern of AIS trajectories. The proposed model is consists of training the pattern, clustering, and classification. We use a machine learning algorithm for training the trajectory patterns. The unsupervised machine learning algorithm, called autoencoder, generates space of trajectory patterns from AIS data. And then clustering the navigational status in this space. The algorithm of clustering runs clustering task recursively until finding optimal groups using pre-classified sample dataset. Finally, the classification module adopts the k-NN (k-Nearest Neighbor) algorithm and can classify the navigational status based on the clustering results. The concept of the proposed model is as shown in Fig. 2.

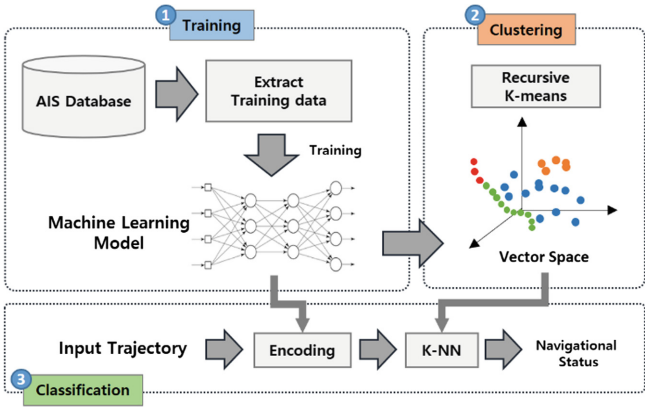


Fig. 2. Concept of proposed model

3.2 Train Traffic Patterns

The training dataset consists of trajectory data received the AIS. The AIS message includes ships' position, course, speed, etc. Also, the training dataset excludes status information in AIS because it contains lots of errors. Therefore, we use the autoencoder which is an unsupervised machine learning model to train the trajectory patterns. The autoencoder model is based on neural network and uses a training data without labels. Also, it is possible to generate a low-dimensional vector space that compresses the training data well. This model sets the target values to be equal to the inputs. And it makes a simpler representation of data by mapping the training data into a low-dimensional space. In this vector space, data with semantically similar characteristics are placed closer together. It is possible to classify navigational status using this feature. The proposed model uses the training set with 49-dimension that extracted from each ship's AIS trajectories as shown in Eq. (1). The training model is a neural network with fully connected that have input, output, and 5 hidden layers. After the training process,

the model generated 10-dimensional vector space. The training set consists of trajectories for 3 min. And the trajectories interpolated linearly with AIS data every 30 s.

$$In_{train} = \{type, Tr_{t0}, Tr_{t1}, Tr_{t2}, Tr_{t3}, Tr_{t4}, Tr_{t5}\} \tag{1}$$

$$In_{train} = Out_{train} \tag{2}$$

$$Tr_t = \{La_t, Lo_t, S_t, C_t, \Delta La_t, \Delta S_t, \Delta C_t\} \tag{3}$$

$$\begin{aligned} where, type &: ship\ type, \\ Tr &: trajectory, \\ La &: latitude, \\ Lo &: longitude, \\ S &: speed\ of\ ground, \\ C &: course\ of\ ground \end{aligned}$$

3.3 Recursive Clustering

The next step of the training model is re-project the training dataset to the vector space and clustering the data with similar characteristics. The *k*-means clustering is a method that is popular for cluster analysis and groups data by minimizing Euclidean distances between them.

In this paper, we proposed recursive clustering algorithm in order to find optimal groups of navigational status. This algorithm based on based on *k*-means and run clustering tasks recursively using the pre-classified dataset. If the cluster includes more than the single status group, the cluster will be split. And, if one status group is contained in several clusters, the cluster will be merged (Fig. 3).

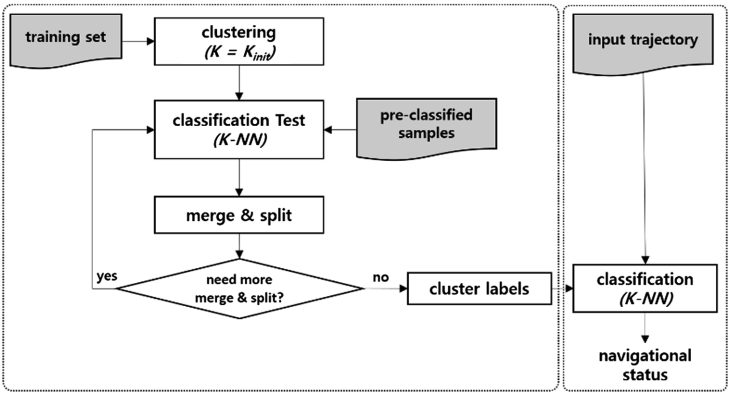
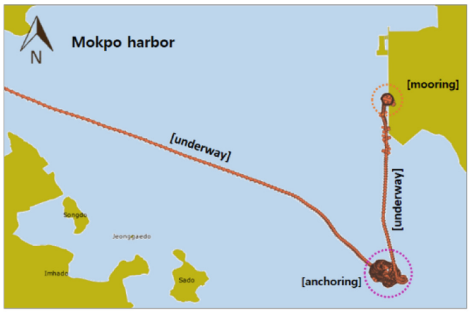


Fig. 3. Flow chart of recursive clustering

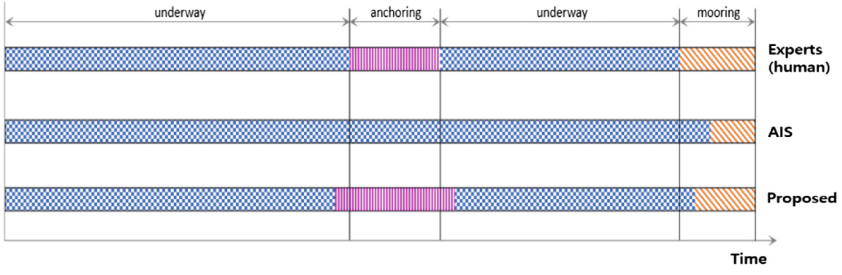
If there no more merge or split in the clusters, the clustering task is ended. In the classification step, new input trajectory data is projected into the vector space and then classify the navigational status. In this paper, we showed an example using *k*-NN (*k*-nearest neighbors) algorithm for classification. However, other methods with good classification performance can be able to apply using the clustering results. For example, SVM (Support Vector Machine) or deep learning.

4 Experiments

We collected AIS trajectory data on the Mokpo harbor for one month and it is 1000 million cases. The training data was extracted from this dataset, and train the machine learning model using this training data. The result of the training, we generated 10-dimensional vector space that encoded trajectory. The training dataset is projected into the vector space, and then runs recursive clustering. The initial number of cluster (*k)* is 30. The sample data of 100 cases are selected by navigational status and classified the status in advance. The test dataset is selected concatenated trajectory data for the single ship from the dataset that is independent of the training dataset.



(a)



(b)

Fig. 4. Experimental result, (a) trajectory overview, (b) classification results

We reviewed the classification results from the expert group, AIS, and the proposed model. In the result of the expert group, the test trajectory consist of three status, "underway", "anchoring", and "mooring". However, there is no "anchoring" status in this trajectory because the AIS information is not updated.

Figure 4 shows the experimental results. Figure 4(a) is an overview of the trajectory that the ship is entering port via the anchorage area. Figure 4(b) shows the classification results of navigational status. The first row is classification result by the expert group, and the second one shows the navigational status data from AIS. The third line shows the classification result of proposed method, which is similar to the expert's results.

5 Conclusion

In this work, we proposed the automatic classification model for ships' navigational status. We describe how to train trajectory patterns using the semi-supervised machine learning algorithm, and introduce the recursive clustering method for clustering the navigational status. Also, we tested the proposed model with the trajectory data. As a result, the status classification from the proposed model is similar to expert opinion and more accuracy than AIS navigational status. In addition, more test is needed to validate the performance of the proposed model in the actual maritime environment. The proposed model is a basic concept of navigational status classification method using prior knowledge. Also, this model would be able to apply to the various applications of maritime analysis. Finally, it is expected that various machine learning approaches which are under development in recent years can contribute to the maritime safety field.

Acknowledgement. This research was supported by a grand from Endowment Project of "Development of core technology for the analysis and reproduction of maritime accidents through simulations" funded by Korea Research Institute of Ships and Ocean Engineering (PES9350).

References

1. Regulations for carriage of AIS (2015). http://www.imo.org. Accessed 16 Feb 2015
2. Perez, H.M., et al.: Automatic identification systems (AIS) data use in marine vessel emission estimation. In: 18th Annual International Emission Inventory Conference, vol. 14 (2009)
3. Gloaguen, P., et al.: An autoregressive model to describe fishing vessel movement and activity. Environmetrics **26**(1), 17–28 (2015)
4. De Souza, E.N., et al.: Improving fishing pattern detection from satellite AIS using data mining and machine learning. PLoS ONE **11**(7), e0163760 (2016)
5. Pallotta, G., Vespe, M., Bryan, K.: Vessel pattern knowledge discovery from AIS data: a framework for anomaly detection and route prediction. Entropy **15**(6), 2218–2245 (2013)

The Retrieval of Regions with Similar Tendency in Geo-Tagged Dataset

Taehyung Lim, Woosung Choi, Minseok Kim, Taemin Lee, and Soonyoung Jung(✉)

Department of Computer Science and Engineering, Korea University, Seoul, South Korea
{th_lim, ws_choi, rlaalstjr47, taeminlee, jsy}@korea.ac.kr

Abstract. We consider an application scenario where user want to find regions that have similar tendency about a certain issue, e.g., looking for regions that are neutral to new welfare policies. Motivated by this, we present a novel query to retrieve regions with similar tendency, named ρ-Dense Region Query (ρ-DR Query), that returns arbitrary shape of regions whose tendency satisfy the ρ-dense constraint. We design a basic algorithm to find all regions with similar spatial textual density that we define in this paper, and also propose an advanced algorithm that performs more efficiently. We conduct experiments to evaluate the performance of both algorithms, and the experiments prove the advanced algorithm is superior to the basic algorithm.

Keywords: Geo-tagged data · Spatial textual query · Region retrieval

1 Introduction

Increasing use of smart devices and various social network services, huge volumes of geo-tagged data keep being updated and become available on the spatial textual analysis which take into account both user's location and user's thought about a certain issue. If we score the user's tendency about a certain issue, we can compute the regional tendency about the issue and find similarity between different regions.

Previous works usually focused on retrieving the hottest region of interest based on relevance score or the densest region based on cardinality.

We propose the new type of query to retrieve all regions with similar tendency about a certain issue in geo-tagged dataset. We define the concept of the spatial textual density to compute the regional tendency from geo-tagged data. We present a novel algorithm to find the all regions which have similar the spatial textual density.

2 Related Work

In this section, we review related works for the retrieval of region in geo-tagged data set. The problem presented in this paper is related to region search and spatial keyword group query.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 42–47, 2020.
https://doi.org/10.1007/978-981-13-9341-9_8

Several studies [2–4] propose to solve the maximum range sum in spatial databases. The region in these studies are defined as rectangle with fixed length and width. Liu et al. [1] study the problem searching subject-oriented top-k hot regions in spatial databases. They only consider the region connected subgraph by road segment with the length constraint. In real world, the region we often want to find may have an arbitrary shape.

The Reverse Spatial Textual k Nearest Neighbor query [10] focuses on finding objects that take the query object as one of their top-k spatial textual similar objects. Several studies [11, 12, 17] propose the solution for collective spatial keyword query, which finds a group of objects that collectively contain query keywords and that spatially close to a query point. Skovsgaard et al. [5, 6] propose the method to find a group of objects maximizing scoring function considering the distance and textual relevance. Zhang et al. [7, 8] find a set of the most n relevant objects to given m keywords. Lu et al. [10] propose the method for finding clusters minimizing scoring function with regard to the distance and the textual relevance, which is based on DBSCAN. Jinfeng et al. [16] propose the method for finding dense region in spatiotemporal databases.

3 Problem Statement

In this section, we present the relevant definitions used throughout this paper and the formal definition of the problem.

We consider the set of data objects, denoted by D, in 2-dimensional Euclidean space E. each object $o \in D$ is a triple (o_x, o_y, o_w) of an object location, (o_x, o_y) and a score of the inclination of an object to score a textual description, o_w.

Definition 1: *l*-Square. The l-square of p, denoted by S_p^l, is a square centered at p with edge length l. S_p^l is a sub-region of the space E and $p(p_x, p_y)$ is an arbitrary point in the space E. The object set in S_p^l is denoted by O_p^l

Definition 2: Spatial-Textual Density (ST Density). A spatial-textual density of S_p^l is defined as $d(S_p^l)$,

$$d(S_p^l) = \frac{\sum_{o \in O_p^l} o_w}{\left\| O_p^i \right\|},$$

where $\left\| O_p^l \right\|$ is the number of objects resided in S_p^l.

Definition 3: ρ-Dense. Given $\rho > 0$, a S_p^l is **ρ-dense** with regard to l if $|d(S_p^l) - \rho| \leq u$, where u is the allowed bound by user. If any sub-square S_p^l of the region R is ρ-dense, the R is called **ρ-dense region** the with regard to ρ and l.

We now address the ρ-dense region query. Finding a ρ-dense region can be considered as finding all ρ-dense l-square in a given space.

Definition 4: ρ-Dense Region Query (ρ-DR Query). Given S and D, a ρ-DR query takes three arguments < ρ, *l*, *E* > , where ρ is a value of the density, *l* is an edge length constraint of the square, and *E* specifies a region of interest. Let P denote the set of center points of all ρ-dense S_p^l in *E* with regard to ρ and *l*. The query returns the ρ-dense region *R* such that $R = \bigcup_{p \in P} S_p^l$. In other words, any subspace of a ρ-dense region *R*, which is represented by *l*-sized square region, is ρ-dense.

4 Algorithm

We proceed to present the algorithms for the ρ-DR query. We adopt the line sweep algorithm [18] to retrieve the regions which satisfy a given density with bound and the grid index structure [19] for query processing.

4.1 Basic Approach

Given a ρ-DR query, the straightforward solution is to first find all ρ-dense *l*-squares in a given space and then to merge all ρ-dense *l*-squares. The merged regions are returned as the result. Based on this solution, we design the **Basic Algorithm**. Algorithm 1 is the pseudo code of the Basic Algorithm.

To solve the problem, an *l*-square sweeps the space by dimension. It is inefficient since the cost sweeping the entire space is expensive.

It first finds all the candidate objects that might be included in ρ-dense region. Based on the grid cells intersecting with given *E*, if the ST density of all objects included in four adjacent cells, e.g., $c_{i,j}$, $c_{i,j-1}$, $c_{i-1,j}$, and $c_{i-1,j-1}$, is higher than the lower bound of density, ρ-*u*, it is possible that all these objects might be included in ρ-dense region. Function GetCandiate tries to find all the candidate objects. The candidate objects in *o_list* are sorted in ascending order of x value and acts as a center point of *l*-square in the next step. It performs a retrieval based on these objects.

Before the retrieval step, we need to do additional work. In the process of retrieval, the sweeping square intersects with fixed objects. We define the moment for the object to meet the border of the square as an *event*, *e*. There are two kinds of event, one is an object enters *l*-square, and another is an object comes out of *l*-square. Function SetEvent handles this task to every objects needed.

To retrieve an available range of ρ-dense *l*-square, it sweeps the space in the *x*-axis direction based on the candidate object *o*, where the event coverage of *o* is o_y-$l < e_y \leq o_y + l$. It calls the function SetEvent and then issues the **ρ-DenseRange** query whenever event e_x located in event coverage of *o* occurs. The ρ-DenseRange query returns the ρ-dense range of the *y*-axis direction when $x = e_x$. Whether the retrieved range is empty or not, it adds to ρ-*dr_list*. The union of all range in ρ-*dr_list* is final result.

Algorithm 1. Basic Algorithm(Query q(ρ, l, E), Cell c, u)

ρ-$dr\_list \leftarrow \phi$
for each cell $c_{i,j}$ do
    $o\_list \leftarrow$ GetCandidate($c_{i,j}$, ρ, u)
/* search ρ-Dense Region */
for each object o in o_list do
    $x\_list \leftarrow$ SetEvent(o, ρ, u, l, c)
    /* sweep in x-axis direction */
    for each event e_x in x_list do
        ρ-$dr\_list \leftarrow \rho$-$dr\_list \cup \rho$-DenseRange(e_x, ρ, u, c)
return ρ-dr_list

When ρ-DenseRange query issues, function SetEvent is called to set objects to events in the y-axis direction. it sweeps the space in the y-axis direction based on the event e_y, where the event is located in o_x-$l/2 < e_x \leq o_x + l/2$. Whenever event occurs, the ST-density of S_p^l is calculated. If ST-density of S_p^l satisfies the ρ-dense condition, function yRange computes the ρ-dense range in the y-axis direction. More than one ρ-dense range can be retrieved for each event, but only ranges intersecting with the event coverage of o can be included in the result.

Algorithm 2. ρ-DenseRange(e_x, ρ, u, l, c)

y_range $\leftarrow \phi$
$y\_list \leftarrow$ SetEvent(e_x, ρ, u, $l/2$, c)
    /* sweep in y-axis direction */
    for each event e_y in y_list do
        if ST-density $|d(S_p^l) - \rho| \leq u$ then
            temp $\leftarrow y$Range(e_y, ρ, u, c)
            if temp $\cap$ event coverage of o then
                y_range $\leftarrow$ y_range $\cup$ temp
return y_range

4.2 Advanced Approach

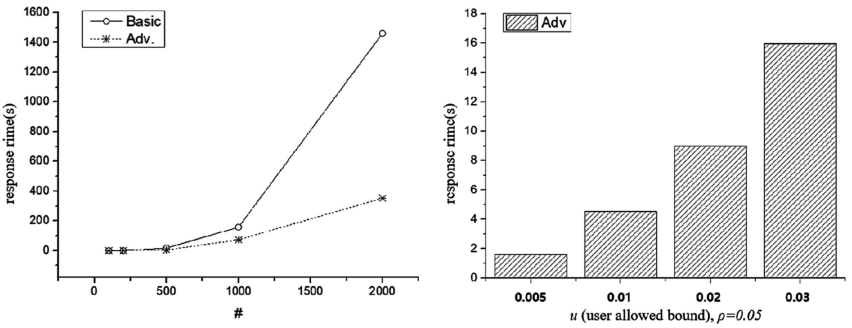
The basic approach is inefficient due to the following reasons. It sweeps for every candidate object. Although the cost of a line sweeping the entire space is basically expensive, it sometimes performs unnecessary sweeping the meaningless space and duplicate sweeping the same space.

We propose the advanced approach to solve these disadvantage and to outperform the basic approach. We design the **Dual-Lane algorithm** to eliminate unnecessary sweeping the space, where any part of the swept space cannot be included in the final result. The method to sweep the space is basically same as Basic Algorithm. Main difference is to sweep the space by the upper and lower lanes separately relative to a candidate object o in the x-axis direction. The event coverage of the upper lane is

$o_y < e_y \leq o_y + l$ and the event coverage of the lower lane is $o_y - l < e_y \leq o_y$. When in the upper lane sweeping in the x-axis direction an event satisfies the ST density condition, it only sweeps the space above the candidate object o in the y-axis direction. When in the lower lane sweeping, it only sweeps the space below the candidate object o in the y-axis direction. The combination of both the upper and the lower result is the final result.

5 Experiment Evaluation

In order to evaluate our proposals, we conduct the experiment to study the performance of the proposed algorithms: the basic approach and the advanced approach. We generate synthetic data sets of size 100, 200, 500, 1K, and 2K. We set the search space size of 2000 × 2000. We evaluate the performance of the Basic Algorithm and the Dual-Lane Algorithm under the different parameter settings. We vary edge length l, ST density ρ, and user allowed error bound u. But due to the lack of space we omit several experiment results. All algorithms were implemented in Python 3.6 on Window 10, and run on Intel(R) Core(TM) i5-4670 CPU @3.40 GHz with 16.0 GB RAM.



Acknowledgments. This research was supported by the Korean MSIT(Ministry of Science and ICT), under the National Program for Excellence in SW(2015-0-00936) supervised by the IITP (Institute for Information & communications Technology Promotion).

References

1. Liu, J., Yu, G., Sun, H.: Subject-oriented top-k hot region queries in spatial dataset. In: Proceedings of the 20th ACM International Conference on Information and Knowledge Management, pp. 2409–2412. ACM (2011)
2. Tao, Y., Hu, X., Choi, D.W., Chung, C.W.: Approximate MaxRS in spatial databases. Proc. VLDB Endowment **6**(13), 1546–1557 (2013)
3. Choi, D.W., Chung, C.W., Tao, Y.: A scalable algorithm for maximizing range sum in spatial databases. Proc. VLDB Endowment **5**(11), 1088–1099 (2012)

4. Cao, X., Cong, G., Jensen, C.S., Yiu, M.L.: Retrieving regions of interest for user exploration. Proc. VLDB Endowment **7**(9), 733–744 (2014)
5. Bøgh, K.S., Skovsgaard, A., Jensen, C.S.: GroupFinder: a new approach to top-k point-of-interest group retrieval. Proc. VLDB Endowment **6**(12), 1226–1229 (2013)
6. Skovsgaard, A., Jensen, C.S.: Finding top-k relevant groups of spatial web objects. VLDB J. Int. J. Very Large Data Bases **24**(4), 537–555 (2015)
7. Zhang, D., Chee, Y.M., Mondal, A., Tung, A.K., Kitsuregawa, M.: Keyword search in spatial databases: towards searching by document. In: IEEE 25th International Conference on Data Engineering, 2009, ICDE 2009, pp. 688–699. IEEE (2009)
8. Zhang, D., Ooi, B.C., Tung, A.K.: Locating mapped resources in web 2.0. In: 2010 IEEE 26th International Conference on Data Engineering (ICDE), pp. 521–532. IEEE (2010)
9. Wu, D., Jensen, C.S.: A density-based approach to the retrieval of top-k spatial textual clusters. In: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, pp. 2095–2100. ACM (2016)
10. Lu, J., Lu, Y., Cong, G.: Reverse spatial and textual k nearest neighbor search. In: Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data, pp. 349–360. ACM (2011)
11. Cao, X., Cong, G., Guo, T., Jensen, C.S., Ooi, B.C.: Efficient processing of spatial group keyword queries. ACM Trans. Database Syst. (TODS) **40**(2), 13 (2015)
12. Long, C., Wong, R.C.W., Wang, K., Fu, A.W.C.: Collective spatial keyword queries: a distance owner-driven approach. In: Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data, pp. 689–700. ACM (2013)
13. Wu, D., Yiu, M.L., Jensen, C.S.: Moving spatial keyword queries: Formulation, methods, and analysis. ACM Trans. Database Syst. (TODS) **38**(1), 7 (2013)
14. Wu, D., Yiu, M.L., Jensen, C.S., Cong, G.: Efficient continuously moving top-k spatial keyword query processing. In: 2011 IEEE 27th International Conference on Data Engineering (ICDE), pp. 541–552. IEEE (2011)
15. Bouros, P., Ge, S., Mamoulis, N.: Spatio-textual similarity joins. Proc. VLDB Endowment **6** (1), 1–12 (2012)
16. Ni, J., Ravishankar, C.V.: Pointwise-dense region queries in spatio-temporal databases. In: IEEE 23rd International Conference on Data Engineering, 2007. ICDE 2007, pp. 1066–1075. IEEE (2007)
17. Cao, X., Cong, G., Jensen, C.S., Ooi, B.C.: Collective spatial keyword querying. In: Proceedings of the 2011 ACM SIGMOD International Conference on Management of data, pp. 373–384. ACM (2011)
18. Souvaine, D.: Line Segment Intersection using a Sweep Line Algorithm. Tufts University, Medford (2005)
19. Rigaux, P., Scholl, M., Voisard, A.: Spatial Databases: With Application to GIS. Elsevier, Amsterdam (2001)

A Programmable Big Data App Container Architecture for Big Data as a Service

Euihyun Jung(✉)

Department of Convergence Software, Anyang University, Anyang City, Korea
jung@anyang.ac.kr

Abstract. In various academic and industrial fields, big data has immensely contributed to finding new and unique insights. In the big data technology, Spark occupies an important position, but its use is confined into either the code-driven or the query-driven ways. We suggest an App container architecture that combines the advantages of two traditional ways. The proposed architecture provides more extensible and capable than the query-driven way. At the same time, it also supports easy call mechanism that the code-driven way cannot provide. Moreover, it presents the structure that shows how Big Data as a Service (BDaaS) is organized.

Keywords: Big data · Hadoop · Big Data as a Service · Spark

1 Introduction

Big data has been a cornerstone of the modern IT technology and has been adopted more and more to extract meaningful insights in most academic and industrial fields such as Biology, Healthcare, Manufacturing, etc. [1, 2, 3]. Nobody denies the importance of the big data technology and everybody is willing to adopt it to his/her field in dreaming of big success.

There are several factors that realize the advent of big data technology. First of all, the price of mass storage, network, CPU for data processing has been going down constantly. Also, the data volume is literally explosive and the number of data sources have been increasing. Beside these known factors, experts may have their own explanations of how big data is widely used in short times, but most professionals will agree that Hadoop [4] is the most important trigger for the technology. After it is possible to store and analyze a huge volume of data with cheap hardware and various analysis methods with Hadoop, the big data technology was no longer a hypothesis but became a realistic one.

In these days, Spark [5] has been the major player in the Hadoop ecosystem. Initially, Spark was an alternative one to Hadoop, but it soon took a seat of Hadoop. It speeds up the analysis by using RAM instead of the hard disk and suggests a new and

This research was one of KOREN projects supported by National Information Society Agency (18-951-00-001).

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 48–53, 2020.
https://doi.org/10.1007/978-981-13-9341-9_9

convenient data model named Resilient Distributed Datasets (RDD) [6]. Currently, there are two ways of using Spark. The first way is the code-driven way where programmers make a Spark application with programming languages such as Scala. Another way is the query-driven way where data scientists analyze data by using SQL-like query languages. These two ways have their own pros and cons. The code-driven way provides programmers with the unlimited capability of analyzing methods, but it is not easy for ordinary data scientists to make applications for their purpose. In contrast, the query-driven way enables data scientists to investigate the data easily and concisely, but the way cannot provide customized and advanced analysis due to the limitations of query languages.

In order to tackle the situation, we suggest an App container architecture that combines the advantages of the code-driven and the query-driven way. In the architecture, programmers make Spark apps which data scientists can select and use to perform various analysis via a dashboard UX or REST API call. If a data scientist needs a specific analysis method, he or she just asks programmers to make a Spark app that supports the method.

The proposed architecture is located in the middle of the existing ways. It has better extensibility than the query-driven way, but worse than the code-driven way. However, from the point of usability and learning curve, the case is opposite. Unlike the existing ways, the proposed architecture shows the possibility of Big Data-as-a-Service (BDaaS) [7]. With this architecture, not only programmers but also ordinary users can analyze data easily by utilizing the full functionalities of Spark via BDaaS way.

The rest of this paper is organized as follows. Section 2 describes the design of the proposed app container architecture. Section 3 explains the implementation and evaluation of the architecture and Sect. 4 concludes the paper.

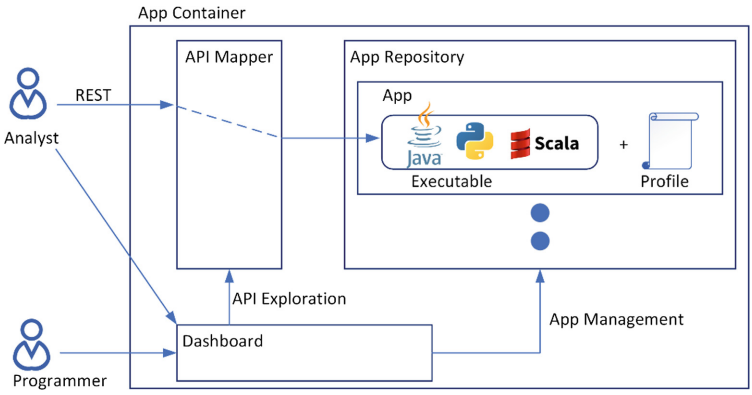


Fig. 1. The structure of App container

2 Design of App Container Architecture

2.1 A Proposed Architecture

As shown in Fig. 1, the proposed architecture consists of *API Mapper*, *App Repository*, and *Dashboard*. There are two roles in using the architecture; programmers and analysts. Programmers usually make an App and upload their App to App container using the *Dashboard*. Each App has an executable which runs on Spark cluster on analysts' requests. It also has a profile that describes a call signature and metadata about the executable. The uploaded App is located on the *App Repository* and it is registered its API signature to the *API Mapper* using its profile.

Analysts can explore the uploaded Apps with the *Dashboard* and request the execution of an App from the *Dashboard* or via a REST API call. When the *API mapper* gets an execution request for an App, it finds the proper App and executes the App on the Spark cluster and returns the results.

2.2 Components of App

Table 1. Metadata in the App profile

Element	Description
uid	A unique id of the App. It is a combination of organization, author name, and app's name
name	The name of the App
version	The version of the App
desc	The human-readable text describing the App
lang	A programming language that makes the App
signature	An API call signature of the App with URI Template [8]
parameters	The parameter lists of the App. Each parameter in the list consists of a name and data type part
callback	A callback method after execution

Each App has two components; the first one is a Spark executable and another one is the profile about the App. The executable can be made with several languages supporting Spark such as Java. The profile contains metadata for the App execution including parameter descriptions. The metadata is summarized in Table 1.

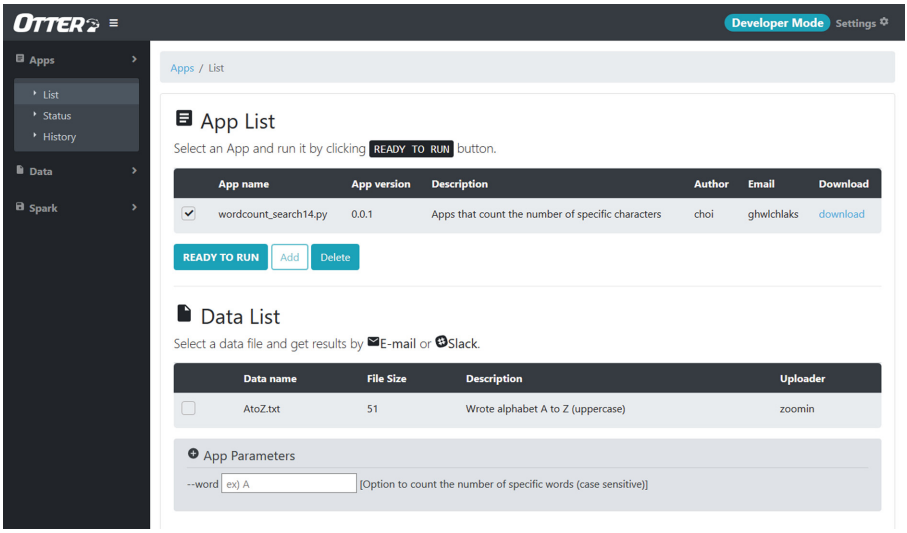


Fig. 2. The execution of an uploaded App via the Dashboard.

2.3 Execution of App

With App Container's *Dashboard*, Analysts can see the uploaded Apps and request an execution of the uploaded Apps shown in Fig. 2. When an analyst selects an App, the data list is automatically displayed. Then, the analyst can choose a data file as the input data of the App. After selecting the target data, the analyst set the parameters of the App and initiates the execution of the App.

The executable's code

```
from pyspark import SparkContext, SparkConf
import argparse
from ayu import noti

parser = argparse.ArgumentParser()
sc = SparkContext()
args = parser.parse_args()
text_file =  sc.textFile(args.src)
counts = text_file.flatMap(lambda line: line.split(" ")) \
          .map(lambda word: (word, 1)) \
          .reduceByKey(lambda a, b: a + b)
output = counts.collect()
if args.mode:
   if "odd"==args.mode:
      print("odd")
      for (word, count) in output:
         if(count %2 == 1):
            noti.report("%s : %i" % (word, count))
   elif "even"==args.mode:
      print("even")
      for (word, count) in output:
         if(count %2 == 0):
            noti.report("%s : %i" % (word, count))
```

The App's profile

```
{
 "uid":"acme/john/count",
 "name":"count",
 "version":1.0,
 "desc":"count letter",
 "lang":"python",
 "signature":"/acme/john/count?mode={mode}&src={src}
              &callback={callback}",
 "parameters":{
  "mode": "int",
  "src": "string"
 }
}
```

The sample of API call

```
/api/v2/count?mode=even&src=hdfs:///choi/line.txt
&callback=email:john@acme.com
```

Fig. 3. The code, profile, and API call of the sample App that counts words in odd or even lines of a given document depending on users' request.

Beside using the *Dashboard*, the proposed system provides analysts with the REST API call for the uploaded Apps. To clarify the process, a sample Spark App is explained in the paper. The App counts words in odd or even lines of a given document depending on users' request. If an analyst wants to count words only in the even line of the document, the analyst requests the execution of the App with the direction argument of "mode" as "even". Since this call signature is written in URI template [8], any kind of argument can be easily described.

In the Fig. 3, the code, the profile, and the call process of the sample App are depicted. In the call, "mode" is set as "even" and "src" is set "hdfs:///choi/line.txt" in the call signature. The call will let the App count words in the even line of the document located in "hdfs:///choi/line.txt".

3 Implementation and Evaluation

Implementation. The proposed system has been implemented with Node.js. The information of the *API Mapper* has stored in MongoDB and the *App Repository* stores Apps in a local storage. The *Dashboard* is made with jQuery and Bootstrap. The App container has been designed to standardize the deployment process and the structure of Apps. Therefore, we expect once developed Apps can be deployed onto and run on any App container with the same manner. It means App containers can simply share Apps so as Docker containers share docker images. We think this structure can be a prominent architecture for BDaaS.

Evaluation. The pros and cons of the App container lie in the middle of the code-driven and the query-driven way. In terms of extensibility and capability, the code-driven way is best because there is no limit for programmers to make an analysis program that utilizes the full functionalities of the Spark system. In contrast, the query-driven way provides great usability and enables data scientists to easily learn how to explore data. The proposed App container takes advantages of both ways and provides users and programmers with the pretty much good level of usability, extensibility, and capability. It also shows the feasibility of how BDaaS is realized. Although BDaaS is expected to be the next generation big data service, there has not been a widely accepted method yet Table 2 represents the comparative evaluation of the proposed system and the existing ways.

Table 2. Evaluation of the App container and the existing ways

Factor	Code-driven	Query-driven	App container
Extensibility	Best	Bad	Good
Capability	Best	Moderate	Good
Usability	Bad	Best	Good
Learning curve	Steep	Shallow	Moderate
Support of BDaaS	Not considered	Not considered	Yes

4 Conclusion

In most academic and industrial fields, big data has played an important role to find meaningful insights. Among the big data technology, Spark is most widely used because of its outperforming speed and built-in modules for streaming, SQL, and machine learning. However, the current usage of Spark is divided into two ways only; the code-driven way and the query-driven way. The code-driven way is powerful but it is difficult for ordinary data scientists not being familiar with programming. Otherwise, the query-driven way is easy and convenient to explore data, but it has a limitation in doing the advanced analysis.

In this paper, we suggest an App container architecture that combines the advantages of the existing two ways. The proposed architecture provides better extensibility and capability than the query-driven way. It also supports easy calls of the preloaded Apps that the code-driven way cannot provide. The architecture not only proposes the better architecture than the existing ways but also shows a new feasible way of BDaaS. Although this system is not designed for this purpose, theoretically, the Apps in the architecture can be transferred and deployed on other App containers like Docker does. Therefore, we expect this architecture can be a prominent architecture for BDaaS and we continue the research to realize it in the further study.

References

1. Chen, M., Mao, S., Liu, Y.: Big data: a survey. Mob. Netw. Appl. **19**(2), 171–209 (2014)
2. Marx, V.: Biology: the big challenges of big data. Nature **498**, 255–260 (2013)
3. Murdoch, T.B., Detsky, A.S.: The inevitable application of big data to health care. JAMA **309** (13), 1351–1352 (2013)
4. Shvachko, K., Kuang, H., Radia, S., Chansler, R.: The Hadoop distributed file system. In: 2010 IEEE 26th symposium Mass storage systems and technologies (MSST), pp. 1–10 (2010)
5. Zaharia, M., Xin, R.S., Wendell, P., Das, T., Armbrust, M., Dave, A., Ghodsi, A.: Apache spark: a unified engine for big data processing. Commun. ACM **59**(11), 56–65 (2016)
6. Zaharia, M., Chowdhury, M., Das, T., Dave, A., Ma, J., McCauley, M., Stoica, I.: Resilient distributed datasets: A fault-tolerant abstraction for in-memory cluster computing. In: Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation, USENIX Association, pp. 1–14 (2012)
7. Zheng, Z., Zhu, J., Lyu, M.R.: Service-generated big data and big data-as-a-service: an overview. In: 2013 IEEE International Congress Big Data (BigData Congress), pp. 403–410 (2013)
8. Gregorio, J., Fielding, R., Hadley, M., Nottingham, M., Orchard, D.: Uri template (No. RFC 6570) (2012)

Deep-Learning-Based Image Tagging for Semantic Image Annotation

Yoonmi Shin[1], Kwangwon Seo[1], Jinhyun Ahn[2], and Dong-Hyuk Im[1(✉)]

[1] Department of Computer Engineering, Hoseo University, Asan, South Korea
sinyoonmil2@gmail.com, rhkddnjsl53@gmail.com, dhim@hoseo.edu

[2] Department of Management Information Systems, Jeju National University, Jeju, South Korea
jha@jejunu.ac.kr

Abstract. Many multimedia images, which are generated in countless ways, must be semantically stored and managed to search for relevant images. Therefore, studies on image annotation have been conducted and the field is developing at a steady pace. In existing approaches, the user inputs the tag for the image directly to obtain the semantic tag. In this study, the tag is input to the image using a convolutional neural network, which can be used for deep learning. Thus, it the hassle of entering tags for images is eliminated. Through deep learning, the user can automatically input the tag for the image, and the entered tags can be expanded to the resource description framework model in Linked Tag. Finally, the user can perform semantic-based image search using the SPARQL query language.

Keywords: Deep learning · Convolutional neural network · Semantic image annotation · RDF

1 Introduction

As the number of social network service (SNS) users with smartphones increases, the importance of communication and promotion through SNSs is increasing day by day. As images for this purpose are increasing, image annotation techniques have been proposed for storing and managing a large number of images [1, 2]. Image annotation expresses image contents by storing semantic information of images together in image contents, and it is a technique that can process a large number of image data effectively.

In particular, Linked Tag [1] uses the ontology to annotate the tag of the image by expanding it to a resource description framework (RDF) triple, and the SPARQL query language is used so that the user can perform a meaning-based image search. However, Linked Tag has a limitation in which the user directly inputs the tag and connects the value of the predicate between the tags [1].

In this study, image processing using deep learning was added to input tags for images automatically. To implement deep-learning, a convolutional neural network (CNN) was used. CNNs are specialized for images and image recognition [8–10].

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 54–59, 2020.
https://doi.org/10.1007/978-981-13-9341-9_10

As shown in Fig. 1, the predicted value of the input image is output through the CNN model. DBPedia automatically fetches the image-related tags with the output predicted value. As a result, tags in the image are imported more efficiently from existing connection tags. Semantic information retrieval of annotated images can be searched using the SPARQL RDF query language as before. Therefore, it is proposed that image annotation can be made more efficient by automatically inputting tags for images using the CNN.

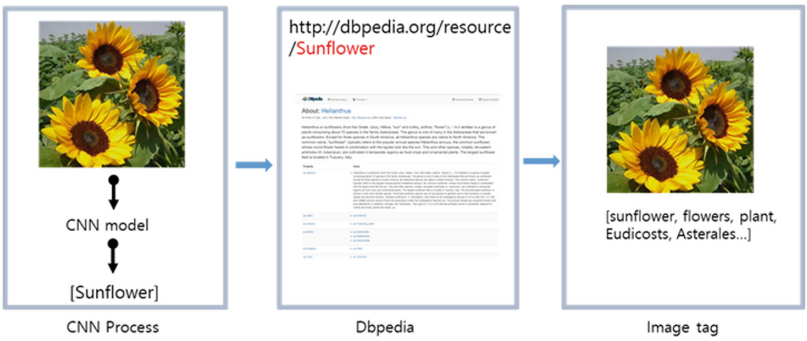


Fig. 1. Automatic image tag process using DBPedia

2 Related Work

Although there were tags on images, there were limits to the use of nouns. To solve this problem, an annotation technique using ontology has been conducted [3–5]. An object relation network (ORN) [3] has been presented as a method to understand and automatically interpret the meaning of annotations on Web images. The entered image recognizes a partial image, called a "segment," and applies it to the probability model to identify the image object. Then, the ontology is used to determine the network of relationships between objects and search for images. Using an ORN enables users to search for images that are similar in meaning to the desired image. A proposed ontology-based image annotation method [4] has the advantage of being able to express meaning in the image. However, there is a disadvantage that the user must establish an ontology in advance. To resolve the problems, an annotation method using the RDF model to expand image tag information was proposed [1]. Because image annotation is performed with the RDF model, semantic-based search for images can be done using SPARQL queries.

CNNs were developed by Yann LeCun and several other researchers in 1998 [6]. However, because of the lack of technology in that period, it did not attract much attention. In 2012 [7], CNNs have begun to attract attention in both computer vision and machine learning, achieving cutting-edge performance for a variety of tasks. The approach [7] is not much different from that proposed in 1998 [6] in its basic composition. A CNN generally comprises a convolution layer and a pooling layer and is performed alternately on two layers. Subsequently performs classification through a

fully connected layer and outputs the results. There is not much change in the structure; however, several methods such as methods, such as rectified linear unit (ReLU) non-linearity, and dropout have been suggested to improve performance and speed.

3 Proposed System

3.1 Deep-Learning-Based Image Tagging

In this research, deep-learning-based image-processing technology was added to existing semantic-based image annotation, and it was expanded to input the tag automatically to the input image instead of the user inputting the tag directly. Deep-learning technology utilizes a CNN [7–10] with excellent performance in image recognition. The CNN model uses TensorFlow[1], an open-source deep-learning library.

As shown in Fig. 2, the CNN is composed of three convolutional layer, two fully connected layer, and a softmax function for classifying five classes.

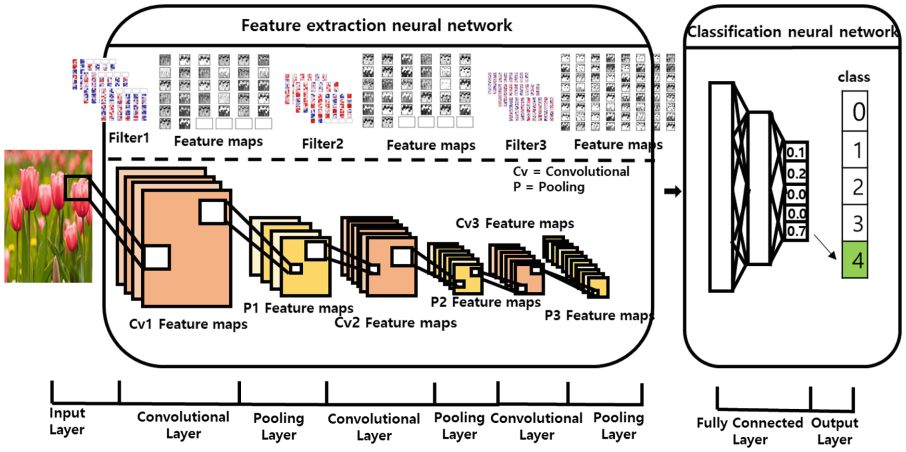


Fig. 2. CNN structure

The first convolution layer forms a 128 × 128 × 32 convolution feature map CV1 through a 3 × 3 × 32 filter of size 128 × 128 × 3 in the input image. Thereafter, image downsizing is performed using a max pooling process to form 64 × 64 × 32 pooling feature maps P1. The result is output as a four-dimensional tensor in the form of '(input image number, feature map vertical size, feature map horizontal size, feature map number)'. If this process is repeated two more times, the final feature map P3 becomes 16 × 16 × 64. Because the classification neural network proceeds after the feature extraction neural network in which the four-dimensional tensor is generated, the four-dimensional tensor is reduced to a two-dimensional tensor. The reduced

[1] http://www.tensorflow.org/.

two-dimensional tensor has 16,384 neurons, and, after two processes in the fully connected layer belonging to the classification neural network, the probability of belonging for each of five classes is estimated through the softmax function. This learning compares the predicted value with the desired output value 20,000 times and reduces the error value to form the optimal model.

When the user inserts the flower image into the learned CNN model, as shown in Fig. 2, it outputs the flower name of the class index having the highest probability among the five classes corresponding to the image. Based on the output flower name, it fetches the tag related to the flower from the information provided by DBPedia and stores it in MySQL with the flower image for use in Linked Tag.

3.2 Semantic-Based Image Annotation Method Using Image Tagging

In this section, a semantic-based image annotation method using Linked Tag [1] is described. The tag for the input image is automatically entered through the CNN. The tags thus obtained automatically link the values of the predicates between tags through the Linked Tag method. Tags associated with linked tags are stored in RDF triple format. When retrieving annotated images stored in RDF, searching is done using SPARQL, the RDF query language.

4 Experiment and System Screen

4.1 Simple Experiment

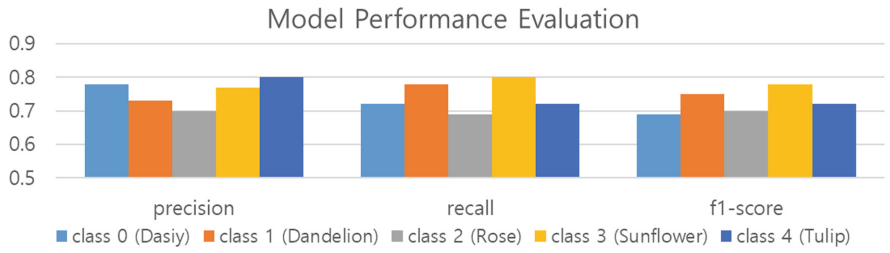


Fig. 3. CNN model performance evaluation

The CNN model returns the index of the most probable class among the five classes through the softmax function to give the predicted value. For example, if the input image is “Daisy,” the index 0 is likely to have the largest value among the five indices through the CNN model. Therefore, when the index 0 has the largest value, “Daisy” is returned to the input image. Figure 3 shows the model performance evaluation graph, which shows how accurately each image predicts each flower image. The learning dataset[2] is labeled with the data for approximately 4000 images and five flower classes.

[2] http://download.tensorflow.org/example_images/flower_photos.tgz.

Approximately 600 images extracted from the learning data to evaluate the learning situation were used for the validation dataset. The validation dataset used here is only used to evaluate the learning situation and does not apply to learning.

Figure 3 shows the precision, recall, and f1-score for the five classes. Precision represents the ratio of what is actually "True" out of "True" predictions. For example, the ratio of the actual "Daisy" image among the predictions of "Daisy" for the input image is shown. Recall represents the ratio of actual "True" predicted to be "True"—for example, when there are 100 images of "Daisy," the ratio of how many of the 100 images accurately predicts "Daisy." The f1-score is the ratio of those correctly predicted when tested using the entire validation dataset. Therefore, the accuracy of the model when viewed through the f1-score is approximately 75%.

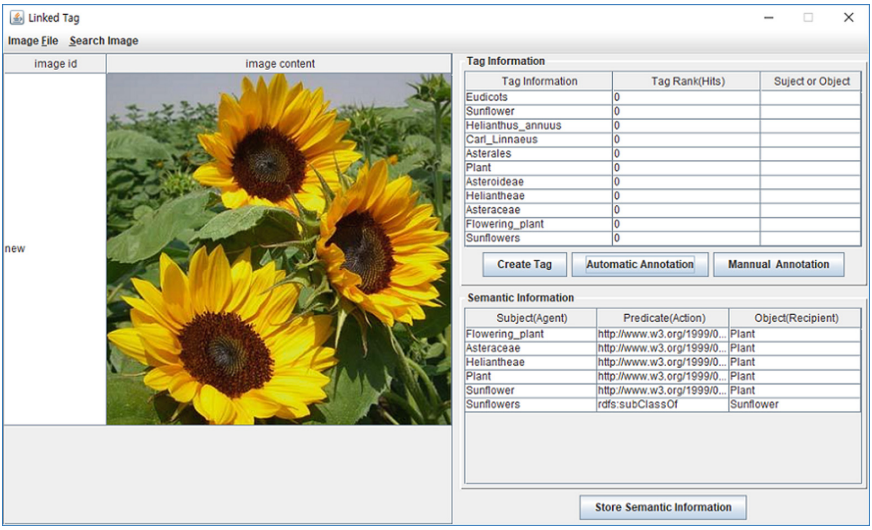


Fig. 4. User interface that shows tags automatically input through CNN and tags input through Linked Tag

4.2 Overall System Screen

First, the user loads the image and stores the image and image indices in MySQL. After pressing the Create Tag button, the loaded image is used as the input image to the CNN model. Subsequently, the image is predicted. With the prediction results, tags related to the image are taken from DBPedia, and the tag for the image is automatically input. The automatically input tags are semantically based image annotations through Linked Tag. The annotated result is stored in the Jena TDB as RDF triple data and then displayed to the user to confirm the tag information through the user interface, as shown in Fig. 4.

5 Conclusion

CNNs were used to store automatically the tag information related to the image in the image contents in DBPedia. Then, annotation information can be saved as an RDF model using Linked Tag, and semantic-based image search can be performed through a SPARQL query. Because the tag for the image is automatically input using deep learning, it compensates for the drawback of the user having to manually input the tag in the image.

In the future, we intend to use mobile devices to enter images in real-time so that image annotation can be possible through deep learning. Furthermore, we plan to design the predicate information tag that links the tags to the image semantically using machine learning to enable the linking of predicate information between tags without using predicate information from DBPedia.

Acknowledgement. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2017R1C1B1003600) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2018R1D1A1B07048380).

References

1. Im, D.H., Park, G.D.: Linked tag: image annotation using semantic relationships between image tags. Multimed. Tools Appl. **74**(7), 2273–2287 (2015)
2. Im, D.H., Park, G.D.: STAG: semantic image annotation using relationships between tags. In: Information Science and Applications (ICISA), pp. 1–2, June 2013
3. Chen, N., Zhou, Q., Prasnna, V.: Understanding web images by object relation network. In: Proceeding of the International Conference on World Wide Web, pp. 291–300, April 2012
4. Hollink, L., Schreiber, G., Wielemaker, H., Wielinga, B.: Semantic annotation of image collections. In: Proceeding of Knowledge Markup and Semantic Annotation Workshop, pp. 41–48 (2003)
5. Park, K.W., Jeong, J.S., Lee, D.H.: OLYBIA: ontology-based automatic image annotation system using semantic inference rules. In: International Conference on Database System for Advanced Applications DASFAA 2007: Advances in Databases: Concepts, Systems and Applications, pp. 485–496 (2007)
6. LeCun, Y., et al.: Gradient-based learning applied to document recognition. Proc. IEEE **86** (11), 2278–2324 (1998)
7. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: NIPS, vol. 1, pp. 1097–1105, December 2012
8. Howard, A.G.: Some improvements on deep convolutional neural network based image classification. In: Proceedings of ICLR, arXiv:1312.5402 (2014)
9. Cireşan, D., Meier, U., Schmidhuber, J.: Multi-column deep neural networks for image classification. In: Proceedings of CVPR, arXiv:1202.2745 (2012)
10. Ciresan, D.C., Meier, U., Masci, J., Gambardella, L.M., Schmidhuber, J.: Flexible, high performance convolutional neural networks for image classification. In: International Joint Conference on Artificial Intelligence, pp. 1237–1242 (2011)

Dynamic Projection Mapping Using Kinect-Based Skeleton Tracking

Sang-Joon Kim[1] and Yoo-Joo Choi[1,2(✉)]

[1] Department of Newmedia, Seoul Media Institute of Technology, 661 Deunchon-dong, Gangseo-gu, Seoul, Korea
gogo59ll@naver.com, yjchoi@smit.ac.kr
[2] Immersive Media Lab., Seoul Media Institute of Technology, 661 Deunchon-dong, Gangseo-gu, Seoul, Korea

Abstract. This paper presents a novel dynamic projection mapping which tracks a skeleton of a user and projects a sequence of images on the body of the user based on robust camera-project calibration without depending on accurate camera calibration using RGB-D image of the Kinect. As the first step of the proposed method, the camera-projector calibration using RGB-D images of the Kinect is performed without the printed calibration board. In the second step, the planes for the torso and palms are defined in 3D world coordinates based on the 2D skeleton information and RGB-D image. In the last step, the projection screen coordinates for the planes of the torso and palms are computed based on the homography between the camera image and the screen image for the projection with considering depth values of the planes of torso and palms. We proved the accuracy and convenience of the proposed method through the projection mapping experiments on a user in various poses.

Keywords: Dynamic projection mapping · Camera-projector calibration · Kinect projection mapping

1 Introduction

Recently, augmented reality (AR) technologies have evolved rapidly and various hardware related to AR have been introduced. Azuma [1] defined AR as systems that have three characteristics: (1) Combines real and virtual, (2) Interactive in real time, (3) Registered in 3-D. Unlike AR using HMD (Head-Mounted Display) and mobile devices, spatial augmented reality is a technology that enhances visual information by projecting images onto the surfaces of real 3D objects and spaces [2]. Spatial augmented reality called projection mapping does not require users to wear any devices and augmented virtual information represented by projection mapping can be shared with multiple users simultaneously [3, 4].

As various tracking systems have been introduced, the dynamic projection mapping has been actively studied recently [3–6]. [5] introduced the dynamic projection mapping onto the body regions, however, this methods did not present the camera-projector calibration method. [6] presented the dynamic projection mapping using the robot arms and Lee et al. [4] presented dynamic projection mapping onto flexible dynamic objects

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 60–66, 2020.
https://doi.org/10.1007/978-981-13-9341-9_11

using real-time image masking method. However, the method of [4] applied to the user who locates in the same position without detail explanation for camera-projector calibration. For the vision-based dynamic projection mapping, the camera-projector calibration [7] is basically necessary. However, a very cumbersome calibration task is required, and the accuracy of camera-projector calibration tightly depends on the accuracy of camera calibration.

In this paper, we propose a novel camera-projector calibration that does not require accurate camera calibration and the printed calibration board. The proposed method uses RGB-D images and computes homographies for each depth level. The screen coordinates for each human body joint points are computed based on the proper camera-projector homography selected by the depth values of each joint points. The experiments using various poses of the user present the accurate dynamic projection mapping results.

2 Camera-Projector Calibration

For camera-projector calibration, we find correspondences between pixels of the projected screen image and camera image points by projecting the chessboard image on the moving board and capturing the projected images using a camera as shown in Fig. 1.

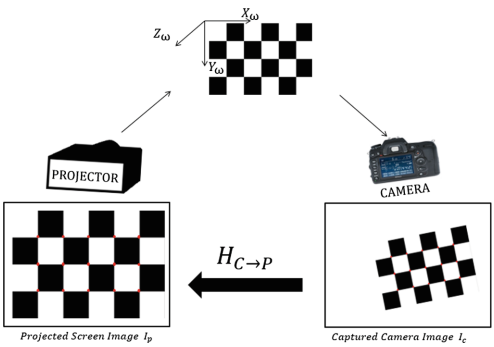


Fig. 1. Homography between the captured camera image and projected screen image.

Equation (1) represents the homography between the projected screen image and the captured camera image.

$$\begin{bmatrix} sx' \\ sy' \\ s \end{bmatrix} = H_{C\to P} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{1}$$

where [x, y] and [x′, y′] mean the pixel coordinates of the captured camera image and the projected screen image, respectively. The chessboard image as shown in Fig. 2 are

used to get the matching point set. The location of the chessboard on the monitor display and four vertices of the chessboard rectangle are adjusted manually to order to project the chessboard onto the moving board properly. To compute homography $H_{C \to P}$, the screen coordinates [x′, y′] and camera image coordinates [x, y] of the corner points of the chessboard are automatically extracted for each pose. The screen coordinates [x′, y′] of fifty-four corner points of the chessboard are determined based on four vertices (V_0, V_1, V_2, V_3) of the chessboard rectangle as shown in Fig. 2 and Eq. 2. Figure 2 represents screen coordinates of fifty-four corner points and four vertices of the 9 × 6 chessboard, and Eq. 2 explains how to compute screen coordinates [x_{ij}, y_{ij}] of a corner point C_{ij}.

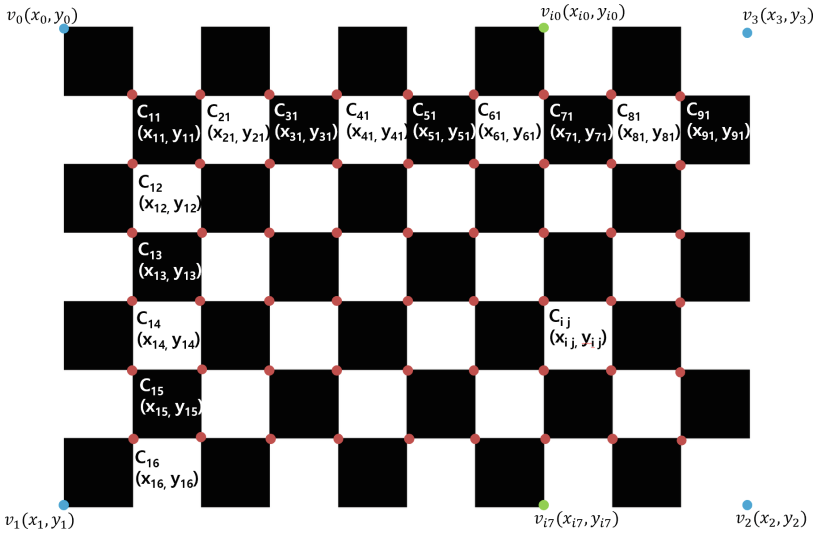


Fig. 2. Screen coordinates of fifty-four corner points (red) and four vertices (blue) of the 9 × 6 chessboard.

$$x_{ij} = x_{i0} * \left(\frac{7-j}{7}\right) + x_{i7} * \left(\frac{j}{7}\right), \; y_{ij} = y_{i0} * \left(\frac{7-j}{7}\right) + y_{i7} * \left(\frac{j}{7}\right), \tag{2}$$

where $x_{i0} = x_0 * (10-i)/10 + x_3 * i/10$, $y_{i0} = x_0 * (10-i)/10 + y_3 * i/10$, $x_{i7} = x_1 * (10-i)/10 + x_2 * i/10$, $y_{i7} = y_1 * (10-i)/10 + y_2 * i/10$.

The camera image coordinates [x, y] of fifty-four corner points of the chessboard are extracted using cv::findChessboardCorners() function of OpenCV library. Figure 3 shows the corner points which are automatically detected in a Kinect RGB image. In order to define the relationship between 3D space, camera image and projection screen image, the white board is moved in the space where the user moves and the chessboard is properly projected onto the whiteboard whenever the location of the white board is changed. Figure 4 shows the different positions of the white board in the same orthogonal distance from the camera. In the different positions of the same depth level

from the camera, the match corner points set of the screen image and the camera image are detected and the homography $H_{C\to P}^{depth}$ for each depth level is defined. *n* homographies are defined for n depth levels as shown in Fig. 5.

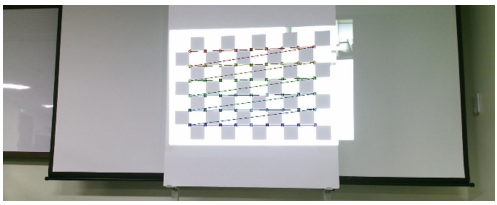


Fig. 3. Fifty-four corner points of the 9 × 6 chessboard extracted in a Kinect RGB image using OpenCV function

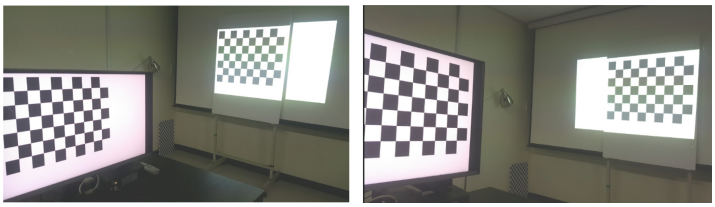


Fig. 4. Projection of the chessboard on the moving board.

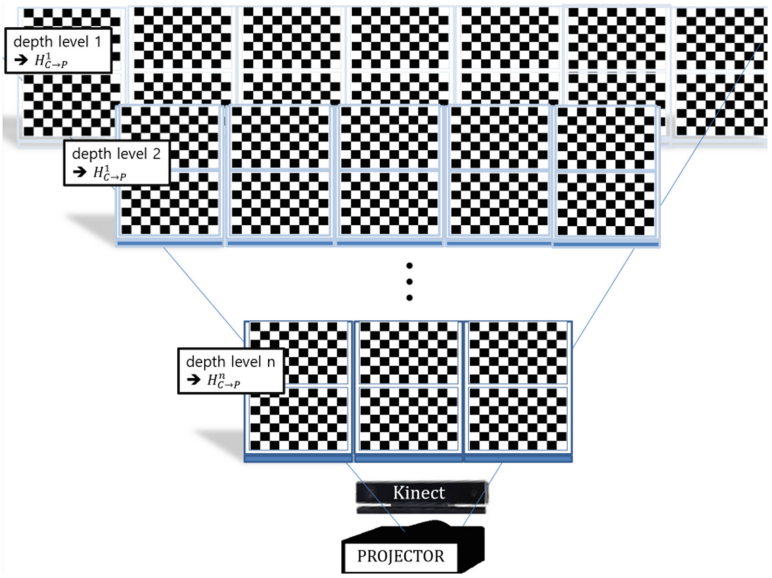
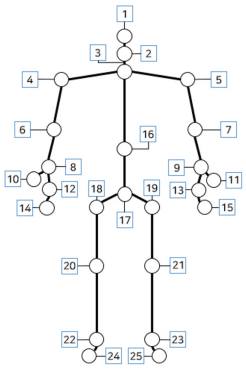


Fig. 5. Different positions of the moving board in the 3D space and homography $H_{C\to P}^{depth}$ for each depth level.

3 Pose Estimation and Projection Mapping

Kinect V2 device includes an RGB color camera with 1920 × 1080, a depth sensor with 512 × 424 and an infrared sensor with 512 × 424. Using Kinect V2, we can track 25 human body joints. Body tracking of Kinect is performed using the depth sensor. So the coordinates of the body joints are aligned with the depth frame. In Sect. 2, the homography for each depth level was computed based on the color image. Therefore, the coordinates of the body joint points should be aligned with the color image. To do this, we use a handy utility named CoordinateMapper which identifies a body joint point from the 3D space to a corresponding point in the color 2D space or in the depth image space. Using CoordinateMapper, the depth value and the color image coordinates of each joint point are acquired in each frame. Figure 6 shows human body joints to be tracked by Kinect V2. The screen coordinate mapping to each joint point can be computed by Eq. 1. At this time, a proper homography among $H^1_{C\to P}..H^n_{C\to P}$ is selected based on the depth value of the joint point. By applying the selected homograpy to Eq. 1, the screen coordinates for projection on the human body joint are acquired.



No.	Item	No.	Item
1	Head	14	HandTipLeft
2	Neck	15	HandTipRight
3	SpineShoulder	16	SpinMid
4	SholderLeft	17	SpineBase
5	SholderRight	18	HipLeft
6	ElbowLeft	19	HipRight
7	ElbowRight	20	KneeLeft
8	WristLeft	21	KneeRight
9	WristRight	22	AnkleLeft
10	ThumbLeft	23	AnkleRight
11	ThumbRigh	24	FootLeft
12	HandLeft	25	FootRight
13	HandRight		

Fig. 6. Human body joints to be tracked by Kinect V2.

4 Experiment Results

For the experiment, we installed a Kinect V2 and a projector as shown in Fig. 7 in the front of the moving white board. For the experiment environment with 3 m × 3 m, we computed three homographies for just three depth levels. As the first experiment, we draw three circles on the center of the torso, that is, the mid spine point and two hands of the moving user. Figure 8 shows the results of projection onto the hands and mid spine point of the user in the different depth levels. Figure 9 represents the result the dynamic projection mapping onto the torso rectangle defined by 4, 5, 18 and 19 joint points of Fig. 16.

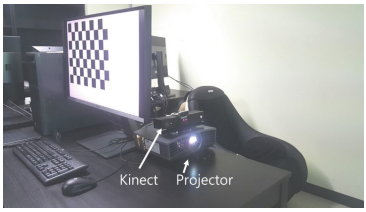


Fig. 7. The Kinect V2 and the projector used in the experiments

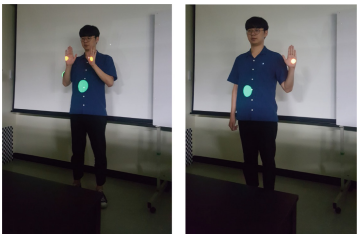


Fig. 8. The results of projection mapping onto the hands and mid spine point of the user in the different depth levels.

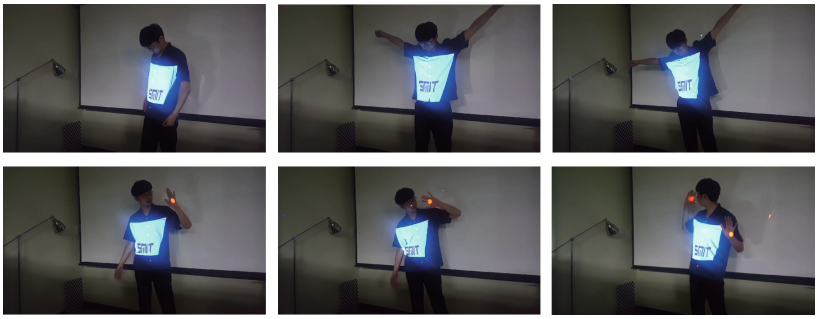


Fig. 9. The results of projection mapping onto the torso rectangle.

5 Conclusion

In this paper, we propose a novel camera-projector calibration method that does not require accurate the camera calibration. The proposed method projects the chessboard onto the moving white board instead of use of the printed calibration board and computes homographies for each depth level. After the camera-projector calibration, it selects the proper the homography based on the depth value of the tracked joint point and computes the screen coordinates for the projection image. The proposed method was implemented using a Kinect V2 and a projector. As a future work, we'd like to develop the dynamic projection mapping onto free deformation objects and to research the methods for the projection delay minimization.

Acknowledgment. This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF-2017R1D1A1B03035718) funded by the Ministry of Education. The human images of this paper were used under publishing consent.

References

1. Azuma, R.T.: A survey of augmented reality. Presence Teleoperators Virtual Environ. **6**(4), 355–385 (1997)
2. Bimber, O., Raska, R.: Spatial Augmented Reality: Merging Real and Virtual Worlds. A.K. Peters, Natick (2005)
3. Narita, G., Watanabe, Y., Ishikawa, M.: Dynamic projection mapping onto deforming non-rigid surface using deformable dot cluster marker. IEEE Trans. Vis. Comput. Graph. **23**(3), 1235–1248 (2017)
4. Lee, J., Kim, Y., Kim, D.: Real-time projection mapping on flexible dynamic objects. In: HCI Korea Conference, pp. 187–190 (2014)
5. Interactive Body Projection Mapping for Hypermetrop. https://vimeo.com/34609484
6. Bot&Dolly, Box. http://bit.ly/19SZdTl
7. Moreno, D., Taubin, G.: Simple, accurate, and robust projector-camera calibration. In: IEEE International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), pp. 464–471 (2012)

Parallel Bidirectional Shortest Path Computation in Graphs Using Relational Database Management Systems (RDBMSs)

Kwangwon Seo[1], MyeongSeok Kwak[1], Yoonmi Shin[1], Jinhyun Ahn[2(✉)], and Dong-Hyuk Im[1]

[1] Division of Computer and Information Engineering, Hoseo University, Sechul-ri, Baebang-eup, Asan-si, Chungcheongnam-do, South Korea
rhkddnjsl53@gmail.com, tech03042@gmail.com, sinyoonmil2@gmail.com, dhim@hoseo.edu

[2] Department of Management Information Systems, Jeju National University, Arail-dong, Jeju-si, Jeju-do, South Korea
jha@jejunu.ac.kr

Abstract. Graph data structures are widely used in computer science fields such as biometric information, navigational systems etc. Recently there has been significant research into quickly calculating the shortest path of a graph using the latest databases such as Neo4j, Spark, etc. Alternatively, the Frontier-Expand-Merge operator (FEM) provides a method to find the shortest path using only SQL in RDBMSs. However, the FEM utilizes sequential searching and iterative aggregate functions to find the shortest path. We propose parallel shortest path searching and table indexing as substitution for the aggregate function. To prove the effectiveness of this approach, we compared each method using experimentation and could demonstrate an improvement of up to 80% in processing speed with our proposal.

1 Introduction

The problem of finding the shortest path on a graph structure has recently been addressed using the latest platforms, various databases and algorithms such as Neo4j [1], Spark [2], bidirectional search [3], etc. The FEM framework [4] had been proposed for the shortest path computation only using RDBMSs and SQL. They attempted to calculate the path using many mature RDBMS functions, which can be largely classified into FEM, bidirectional FEM and Bidirectional Restrictive Breadth First Search (BRBFS) respectively. The FEM is a three-step SQL implementation of the Dijkstra's algorithm [5] consisting of a Frontier-Expand-Merge operator. The FEM is performed in such a manner that the selected node table is continually being updated by an iterative operation which is adding adjacent edges to it. Figure 1 is an explanation of the basic FEM iteration process. Bidirectional FEM uses source to target and target to source processing to achieve progress. Finally, BRBFS is an extension of Breadth First Search (BFS) using partitioned edge table divided by edge weight. However, bidirectional FEM has the limitation that it requires serial order processing so although its

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 67–72, 2020.
https://doi.org/10.1007/978-981-13-9341-9_12

path exploration is bidirectional, it expands its path sequentially and potentially results in bottlenecks.

In this paper, we propose an optimization method which uses parallel multi-threading and table optimization through indexing and SQLs in order to improve the speed of finding shortest path.

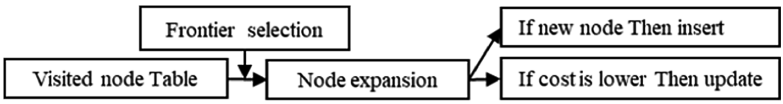


Fig. 1. Basic FEM iteration process with visited table

2 Bidirectional FEM

Bidirectional FEM is an extension of FEM to use two directions. It searches using both forward and backward directions i.e., from source(s) to target(s) and from target(s) to source(s). However, its bidirectional search is not executed in parallel. Its direction just changes to forward or backward based on data in the selected table and therefore it has only one direction at any given time (Fig. 2).

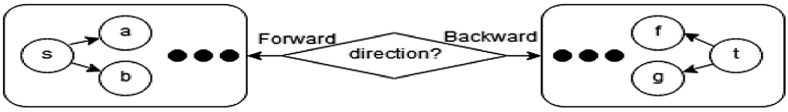


Fig. 2. Bidirectional FEM search process.

3 Bi-thread Path Search

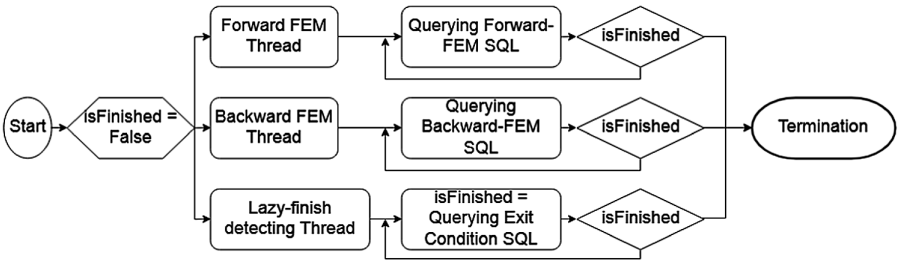


Fig. 3. Main three thread iteration of Bi-thread.

We can use the thread method to find shortest path for each direction simultaneously by separating its process into three threads that consist of forward, backward and terminal detection. Figure 3 shows three threads proceeding independently.

Firstly, we initialize the start and end nodes of each visited table with zero value in distance to source (d2s) attribute. After initialization, forward and backward threads search for the shortest path using SQL, and the discovered node adds its cost, previous node, and visited identifier to the attributes d2s, p2s and ff respectively. Next, two threads find paths until intersection node is discovered. In this process, threads are adding visited node to visited node table and updating variable minCost. Variable minCost is the termination factor of path finding and minCost value is the sum of forward and backward nodes previously visited with minimal d2s value.

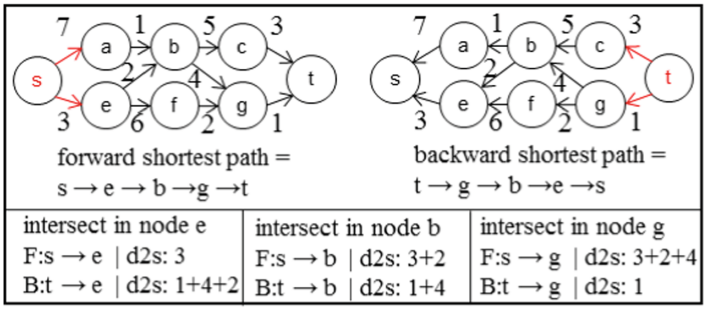


Fig. 4. Termination condition check.

When the intersection node is discovered, minCost is also calculated in the termination condition thread. In the next search, the sum of the newly calculated path's cost exceeds the shortest path cost. Then the threads stop to make both paths connected. Figure 4 shows an example of the termination condition. In Fig. 4, whatever the intersection node is, the variable minCost is 10 and at the next search the d2s value exceeds the minCost and thus the threads are terminated.

After the termination condition has been satisfied, the path connecting process is started in both forward and backward directions. We find the shortest path bidirectionally but the path should be created in the original direction. Therefore, attribute p2s which is in the visited table is mainly used in the process which connects the discovered path with both forward and backward directions.

Forward and backward searching are independent and can proceed regardless of termination conditions. The bottleneck caused by the termination check occurs in calculating variable minCost because the process has to wait until one direction calculation is finished before minCost can be updated. However, there is no further waiting because its process has been independently completed.

4 Visited Table Indexing

In FEM processing, the selecting frontier node operation is continually repetitive from the start until the end costing significant elapsed time. In the bidirectional FEM, the aggregate SQL function Min() is used to select the next frontier with the smallest d2s attribute. However it is not necessary to use the aggregate function to deduct a node (Fig. 5).

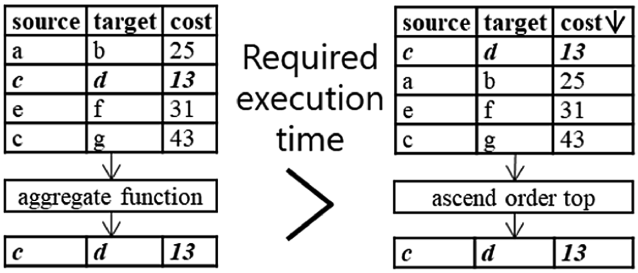


Fig. 5. Required time cost about next frontier node selection.

If the table schema is maintained with ascending sorted order through a btree [6] index, the aggregate function can be substituted with just "select top" SQL, which has low-cost execution time. If the process accesses the visited node with indexed order, the overhead cost of the aggregate function execution can be avoided.

5 Experiments

We implemented JDK10 and evaluated performance by Intel i5-4670 3.4 GHz processors with 8 GB RAM and with the operating system Ubuntu 18.04 LTS. The DBMS was PostgreSQL 10. The target dataset was a 9th DIMACS Implementation challenge [7] undirected weighted graph (Table 1).

Table 1. List of experimentation dataset.

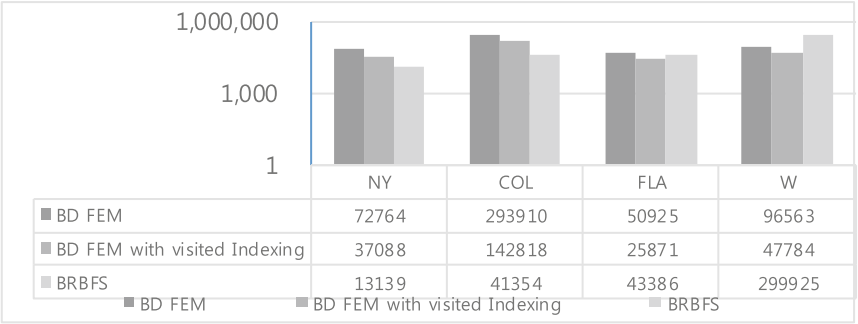
Dataset	Nodes	Edges
New York (NY)	264,346	733,846
Colorado (COL)	435,666	1,057,066
Florida (FLA)	1,070,376	2,712,798
Western USA (W)	6,262,104	15,248,146

The graph of each dataset used in the experiment was named U.S.A. road dataset. Those graphs consisted of undirected weighted edges. Each edge was represented by set of [start node, end node, edge cost]. The results of tests were average value of time-cost in millisecond units and we compared each approach sequentially.

5.1 Comparison with Visited Table Indexing

Then we compared to the other path search algorithm. In Sect. 4, we can replace the aggregate function with the indexing method thereby avoiding the aggregate function. Figure 7 shows the experiment results for each dataset with log-scale using BD as the abbreviation of bidirectional. Figure 7 indicates an approximate 50% performance improvement over normal bidirectional FEM when using indexing. We believe this to

be a significant advantage. We also compared to BRBFS, that is, Bidirectional Restrictive Breadth First Search, which is restrictively optimized BFS with partitioned edge table. BD FEM with visited indexing was worse than BRBFS in a small dataset. However using a larger target dataset BRBFS took far more time than BD FEM because BRBFS has to access its partitioning table many times since that is proportional to the dataset (Fig. 6).

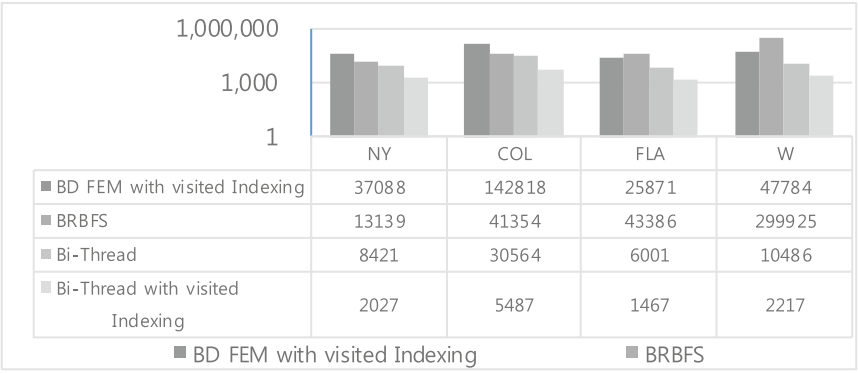


	NY	COL	FLA	W
BD FEM	72764	293910	50925	96563
BD FEM with visited Indexing	37088	142818	25871	47784
BRBFS	13139	41354	43386	299925

Fig. 6. Comparison with visited table indexing

5.2 Comparison with Bi-thread

We observed in Fig. 7 that BRBFS was mostly faster than BD FEM. However, we also saw that in some cases BRBFS was even worse than basic FEM.



	NY	COL	FLA	W
BD FEM with visited Indexing	37088	142818	25871	47784
BRBFS	13139	41354	43386	299925
Bi-Thread	8421	30564	6001	10486
Bi-Thread with visited Indexing	2027	5487	1467	2217

Fig. 7. Comparison with Bi-thread

We wrote each value in the table because some values were excessively high. In general regardless of the table type the bidirectional thread process was faster than the others. Moreover indexing had a positive effect on processing times.

6 Conclusion

This paper proposed improving existing FEM framework through visited table indexing and parallel processing. Firstly, the aggregate function overhead is a time-consuming operation but we can replace that operation with "select top" by indexing. Secondly, the FEM can find the shortest path using the bidirectional method but its processing method is limited by sequential processing. However bi-thread takes an approach that is parallel direction independent. Consequently its thread calculates its own direction's path search instead of its sequential processing by direction. Therefore, the entire processing time is much lower.

Acknowledgement. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2017R1C1B1003600) and Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (No. R0113-15-0005, Development of a Unified Data Engineering Technology for Largescale Transaction Processing and Real-Time Complex Analytics) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. NRF-2018R1D1A1B07048380).

References

1. Neo4j. https://neo4j.com/
2. Apache Spark. https://spark.apache.org/
3. Wagner, D., Willhalm, T.: Speed-up techniques for shortest-path computations. In: Thomas, W., Weil, P. (eds.) STACS 2007. LNCS, vol. 4393. Springer, Berlin (2007)
4. Gao, J., Zhou, J., Yu, J.X., Wang, T.: Shortest path computing in relational DBMSs. IEEE Trans. Knowl. Data Eng. **26**(4), 997–1011 (2014)
5. Dijkstra, E.W.: A note on two problems in connexion with graphs. Numer. Math. **1**, 269–271 (1959)
6. Comer, D.: Ubiquitous B-Tree. ACM Comput. Surv. (CSUR) **11**(2), 121–137 (1979)
7. The 9th DIMACS Implementation Challenge - Shortest Paths. http://www.dis.uniroma1.it/~challenge9/papers.shtml

Authoring Tool for Generating Multiple Experiences of 360° Virtual Reality

Syed Hammad Hussain Shah and Jong Weon Lee(✉)

Department of Software Convergence,
Sejong University, Seoul, Republic of Korea
hammad.shah38@gmail.com, jwlee@sejong.ac.kr

Abstract. 360° video is a special form of digital content which provides visual details of 360° world around the viewer. While watching such video using head mounted display (HMD), it immerses the viewer with experience of a different place. This paper introduces a novel authoring tool which enables the viewers to create multiple interesting experiences of a 360° video. Viewers can read descriptions of experiences provided by their creators and choose anyone to watch. This tool assists users to efficiently follow the selected experiences while watching 360° video. Major challenge in existing authoring tools is that they are difficult to learn. We have made authoring easy for user to create experiences of 360° video by simply watching in virtual reality by wearing an HMD. Another problem in previous authoring tools is that it is difficult to author multiple video experiences of a 360° video. This authoring tool has made it easy.

Keywords: 360° video experience · Authoring tool · Virtual reality

1 Introduction

Virtual reality (VR) has opened new doors for researchers. It has wide range of applications in various fields like education, entertainment, medical, etc. 360° videos have gained more popularity especially in VR due to its high immersiveness [1]. Head mounted display (HMD) makes 360° video content very engaging because viewer has lot to explore in surroundings within virtual world [1]. Due to its increasing popularity; along with the providers of 360° video content, social media platforms like Youtube and Facebook have also enabled users to upload 360° videos [2].

There are multiple challenges in watching 360° video using HMD. Viewer has complete freedom to look around in a 360° video while HMD has limited field of view (FOV). Hence, making it difficult to direct viewers' attention towards the content of their interest. Many authoring tools enable creators of 360° video to direct viewers' attention towards the regions of interest (ROI). A limitation in these authoring tools is that creating multiple experiences of a 360° video is not easy. They are mostly web or desktop based [3, 4]. There can be distinct visual information at different regions of 360° video around user. On the basis of interest users may attend selective regions while watching 360° video. So, experiences of users would vary based on watched ROIs. Our proposed solution is to generate multiple experiences of 360° video in VR by saving users' head orientation after each cycle of a specific time interval while watching the

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 73–78, 2020.
https://doi.org/10.1007/978-981-13-9341-9_13

video. Head orientation determines where user is looking in 360º environment. Another challenge in watching 360° video is searching ROI. Due to large angular distance of head pose from ROI, viewers may not attend it in time which may results into loss of major information. Our focus assistance technique helps viewers in attending ROIs on right time based on generated experiences of 360° video. Our proposed method overcomes the problem of large angular distance between users' current and target FOV. This paper is organized as; Sect. 2 presents previous work related to our study. Section 3 describes the proposed system. Section 4 explains user study with experiment design and results' analysis. Finally, Sect. 5 brings it to end with conclusive remarks.

2 Related Works

Market's influence towards 360° video content generation is seeking much attention of researcher community of VR. As a result, it yields a number of publications about how to make watching experience of 360° video better. Many challenges and their solutions including the problems of finding ROIs and how efficiently following the 360° video's narrative have been discussed by the researchers [5, 6].

In May 2017, Yen-Chen Lin et al. investigated two focus assisting techniques for 360° videos [2]. They studied effect of auto-focus (AF) assistance and visual guidance (VG) techniques on watching 360° videos of sports and tours. According to this study, AF assistance outperformed VG in providing better experience for watching sports video and vice versa in case of 360° VR tours.

In October 2017, another study on assistance technique for finding the out-of-sight ROIs in a 360° video was performed [5]. In this publication, Yung-Ta Lin et al. have proposed the solution of displaying out-of-sight ROIs in video by augmenting the current FOV with the picture of out-of-sight region. Another study about viewing experiences of 360° videos was performed in October 2017 [7]. It was based on viewport's characteristics and their influence on viewing experience. This study presented that moving viewport requires more exploration from viewer than static viewport which requires user's more effort. Moreover, Facebook has also provided guide to 360° video content providers [8]. It allows providers to mark specific points of interests as narrative over the whole course of video.

In June 2016, an authoring tool named as WondaVR was released by Arnaud Dressen et al. [3]. It allows users to create VR experiences by uploading 360° content to system using desktop application. In the same year, another VR technology related company named as InstaVR launched a web-based authoring tool for 360° videos' producers. It was founded by Daniel H. Haga [4]. This tool also helps users to create VR experiences and publish them on any capable device for watching.

3 Proposed Authoring Tool

We have created an authoring tool consisting of two major parts. First part named as the 'Creator Part', records viewer's experience of a 360° video and the other one is 'Viewer Part' for watching these created experiences. For these two parts, we have

created an android 360° video player in unity which renders video on sphere and places viewer at center of it. On top of the player several methods have been implemented for both parts. We have discussed both parts of the authoring tool below.

3.1 Creator Part

During a 360° video in VR, we continuously track user's head orientation using a gyroscope sensor. This determines current viewpoints' positions within 360º rendered scene. Our system saves this head orientation during 360° video after every 100 ms. While watching video user can stop, pause, rewind or forward it. In case, when video:

1. Is paused, viewer's orientation does not get saved until the video starts playing.
2. Rewinds, the system records the experience from the starting time of replay.
3. Is forwarded, viewpoints information for skipped video parts doesn't get saved.

At the end of watching video, user can preview created experience (Discussed in Sect. 3.2). Experience can be edited after previewing it by re-watching the video part which he/she wants to edit. That video part can be reached by using a skip or rewind function of video player. Experience can be saved to the system after previewing and editing it for desired number of times. Before saving experience, user had to describe it by using virtual keyboard and controller. This description assists other viewers in selecting and watching 360° video experience of his/her interest.

3.2 Viewer Part

In this module, we have implemented visual guidance in 360° video player for watching experiences shared by creators. We preferred visual guidance over auto-rotation of viewport for avoiding motion sickness, as discussed by Gugenheimer et al. [6]. Using this system, a viewer can select any one of the created 360° video experiences for watching. Visual indicators guide the viewer towards intended target according to created experience. As an intended target starts becoming clearly visible, visual indicator starts disappearing. In our system, criterion for intended target to be considered as clearly visible is based on two things; approaching 20° right or left and 15° top or bottom to the center of FOV. One major challenge in guiding the viewer according to creator's experience is dynamicity of intended target. During 360° video experience it is continuously needed to find right direction for guiding the viewer. We are doing visual guidance by pointing visual indicator towards right, left, upward, downward, top right corner, top left corner, bottom right corner and bottom left corner as needed.

Another major challenge in transferring the created experience to another user is lack of synchronization with visual indicators. Viewer may not attend the intended target on right time. We have proposed following approaches for this problem:

1. Opacity of the visual indicator increases with the increase in angular distance between the centers of viewer's FOV and the intended target (see Fig. 1b and c). If the current head pose is represented as 'H', head pose to attend the intended target as 'I', angular distance between 'H' and 'I' as 'θ' as shown in Fig. 1a.

(a) (b) (c)

Fig. 1. **(a)** It represents angular distance between current head pose and head pose required for attending intended target. Visual indicator has: **(b)** low opacity due to small angular distance, **(c)** high opacity due to large angular distance

And opacity of visual indicator is denoted by 'O'. We can represent relation between 'O' and 'θ' as in Eq. 1.

$$O \propto \theta \tag{1}$$

2. 180° can be the maximum horizontal angular displacement between two points in 360° world. So, we decrease normal playback rate of video by 25% with each 60° rise in angular displacement. Playback rate will be 25% at 180° (See Eq. 2). When video gets slow down, user can easily assess that he/she is lagging behind the intended target. So, with the help of visual indicator user tries to reach the intended target for bringing video at normal playback rate and continue experience.

$$Playback\ Rate \propto \frac{1}{Angular\ Distance}\ If,\ Angular\ Displacement\ >\ 60^{\circ} \tag{2}$$

4 User Study

A swivel chair which was provided to participants in order to rotate smoothly while watching 360° video in VR. We used a smartphone mounted to Samsung Gear VR connected to a controller. Users had to wear the Gear VR with precision lenses having FOV of 96°. We have used a Samsung Note 8 smartphone with a CPU of Exynos 9 octa-core which can render 360° videos efficiently. This study aimed at the creation and authoring of multiple experiences of 360° video. The participants were provided with two 360° videos of different categories. First video was related to technology evolvement in farming and information about plants' growth inside land. Second video was comprised of virtual tours of more than ten beautiful places of Europe. Viewers had choice to select any one video for watching. Evaluation measures for user study were related to creating and previewing experiences in VR.

4.1 Experiment Design

Interest about visual content varies from person to person. So, we provided choice to participants for selecting the video of their interest. They were informed that their experiences with the 360° video will be recorded. This study is focused on both parts of our authoring tool. In first part; as a creator, user had full freedom of moving his/her head for watching video according to his/her interest for creating experience. In second part, user had to preview his/her created experience as a viewer to assess the accuracy of transferring of experience. While previewing, user had to follow visual guidance as discussed in Sect. 3.2. After previewing experience user had choice to either edit or save the experience after describing it for sharing with others (See Sect. 3.1). All sequential activities of experiment are shown in Fig. 2.

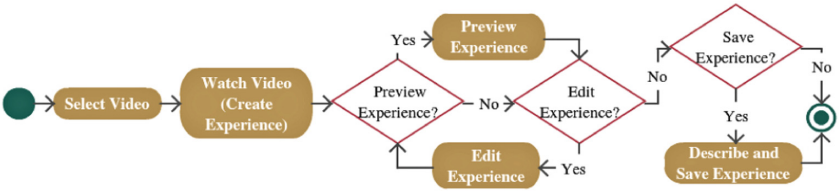


Fig. 2. Activity diagram for creating experiences of 360° video

After saving experience, participants filled a questionnaire. We recruited 25 paid participants aged between 23 and 31 (Mean (M) = 25.4, Standard Deviation (SD) = 1.88). Eight participants had no experience in using VR device, four of the total were frequent users while remaining 13 participants had little experience with VR devices.

4.2 Quantitative Analysis and Results

Designed questionnaire was influenced by the previous research work and simulator sickness questionnaire (SSQ) [9] although SSQ was not strictly followed. We analyzed self-assessed experiences of participants by using mean opinion score (MOS) technique because our main concern was participants' experience with our system. We categorized user test into three categories named as usability, efficiency and satisfaction. Our study was consisted of total nine questions. Score distribution was as; 1 for strongly disagree, 2 for disagree, 3 for neutral, 4 for agree and 5 for strongly agree.

Usability. This measure included four questions and focused on assessing convenience to use the authoring tool and its learnability. 53% of total participants strongly agreed with the fact that proposed authoring tool is very easy to learn and use, 41% just agreed while 6% participants remained neutral about the statement (M = 4.47, S.D = 0.082).

Efficiency. It emphasized on evaluating the effectiveness of transferring of experiences and system's performance. Three questions were asked in this part. 72% participants reported more than 90% similarity between their actual saved experience and when it was guided by our authoring tool. Whereas, 28% participants reported more than 75%

similarity. In case of overall efficiency, 68% participants rated the system as highly efficient, 24% as efficient and 8% scored it as just acceptable (M = 4.6, S.D = 0.12).

Satisfaction. This measure focuses on participants' satisfaction with proposed system. Strong point of our tool is that minimal amount of knowledge is required to use it. Response of participants about proposed system is a good motivation. 72% of total participants were highly satisfied with system's application (Strongly Agreed) while 28% were satisfied (Agreed) up to good level but not very high (M = 4.78, S.D = 0.06).

5 Conclusion and Future Work

We presented an authoring tool to create multiple experiences of 360° video in VR and evaluated its effectiveness. Results of our study depicts that participants felt easy in using our proposed authoring tool due to simple operations as compared to other authoring tools. Watching 360° experience using this tool was reported easy by users due to efficient focus assistance technique and minimum effort required for searching ROIs in video. It increased user's enjoyment because they could easily find interesting information in video through selecting and watching experience created by others. Our future work aims at generation of 360° experiences using computer vision techniques.

Acknowledgment. This work is supported by Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 18CTAP-C142967-01-000000).

References

1. Cummings, J., Bailenson, J.: How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. J. Media Psychol. **19**, 272–309 (2016)
2. Lin, Y.-C., et al.: Tell me where to look: investigating ways for assisting focus in 360° video. In: Proceedings of the ACM 2017 CHI Conference on Human Factors in Computing Systems (2017)
3. Dressen, A., et al.: Wonda VR. https://www.wondavr.com/
4. Haga, D.H., et al.: Insta VR. http://www.instavr.co/
5. Lin, Y.-T., et al.: Outside-in: visualizing out-of-sight regions-of-interest in a 360 video using spatial picture-in-picture previews. In: 30th Annual ACM Symposium on User Interface Software and Technology (2017)
6. Gugenheimer, J., et al.: SwiVRChair: a motorized swivel chair to nudge users' orientation for 360° storytelling in virtual reality. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (2016)
7. Van den Broeck, M., et al.: It's all around you: exploring 360° video viewing experiences on mobile devices. In: Proceedings of the 2017 ACM on Multimedia Conference (2017)
8. New Publisher Tools for Facebook 360° Videos. https://www.facebook.com/facebookmedia/blog/new-publisher-tools-for-360-video
9. Kennedy, R.S., et al.: Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int. J. Aviat. Psychol. **3**, 203–220 (1993)

A Study of Open Big-Data Platform Architecture for Housing Market Analysis

Sang-Hun Lee, Jeong-Ran Yun, Tae-Gyun Kim(✉), and Jung-Min Oh

Korea Land and Housing Corporation, Yuseong-Gu, Daejeon 34047, Korea
{icarus, yjr4l2, raphaelo, ojm}@lh.or.kr

Abstract. There has been increased demand for the development of a system that provides easily accessible open-type housing market information and utilizes that system using large data; the objectives are to increase the positive effects of housing policy and improve understanding of the policy among the people. Since there is the increased necessity of developing a decision support system for housing policy to create reasonable and scientific housing policy, a research on "Development of Housing Market Analysis and Forecasting Model Using Large Data" was started. In this paper, This papers describes the design of the technology equipped with an open platform architecture for internal and external users.

Keywords: Housing market analysis · Open source · Big data platform · Spatio-temporal storage

1 Introduction

To date, housing market forecasts have been conducted through statistical and demand survey. Predicting the housing market is an estimation of demand and supply, which is carried out by many experts and there are various methods. However, although long-term forecasts are possible, there is a limit to how to respond appropriately to short-term demand and supply instantly. In addition, macro analysis and micro analysis must be performed simultaneously in the long-term demand analysis, which also has clear limitations by statistical and demand survey methods.

Therefore, it is necessary to develop a decision - making system using Big Data in order to solve these problems and to make decision of housing policy efficiently. This paper suggest the forecast system to the demand and supply of housing by inputting not only the data of the demand survey but also the raw data of the entire housing market and all actual transaction prices. In addition, the need for a housing big data platform that includes GIS functions to perform spatial analysis in demand and supply forecasting has emerged. This paper designs and builds the big data platform structure using open sources which have scalability and flexibility to prepare for the future increase of data.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 79–84, 2020.
https://doi.org/10.1007/978-981-13-9341-9_14

2 Technology Demand Survey

2.1 Open Platform Software

In a typical big data system, scalability is very important because it deals with large amounts of data. When designing the platform, we should clarify the size and type of data, frequency of processing, and the type of analysis model. However, since there is no case to collect all the data of the housing, the housing market platform is designed based on the open source that can accommodate the data processing capacity and property, variety of the analysis model without being dependent on the vendor. Particularly in the case of analysis programs, users around the world are developing analytical functions, which is very good in terms of speed, accuracy and diversity.

The following techniques have been adapted for the purpose. Technology for collecting, loading and linking: Hadoop, Yarn, HBase, Open SSL, FTP, Crawling technique, Accumulo, Spark, Ambari, Zookeeper. Processing and analysis techniques: Hive, R, Postgre. Technologies for visualization and information delivery: Apache, Tomcat, Q-GIS. Since the main data of the system is structured-data, it is more advantageous to use memory DB when processing big data, but this paper adopted the Hadoop structure for spatio-temporal analysis and future SNS analysis.

2.2 Related System to Be Interconnected

To analyze the housing market, it is essential to collect and organize various data. In order to collect the data in the public area, it is necessary to analyze the Real Estate Transaction Management System (RTMS), the Housing Information System (HIS), the Building Administrative Information System (EAIS), the Real Estate Announcement Price (REAP), the National Spatial Information Portal system (NSIP) [1, 2].

The Building Administrative Information System (EAIS) is an information system that records the whole process of construction administrative civil complaints that manage the construction, housing authorization and permission. In 17 provinces and 228 municipalities, it is used as a standard to manage the most accurate data. However, the architectural administrative information system does not own the information of unauthorized houses because it updates the data in case of accidents and permits.

The Real Estate Transaction Management System (RTMS) is a system that allows the price of real transactions to be registered at the time of the house transaction, and manages the information history about the prices. This is good data for determining the current price of a house, but there is no data that has not been traded for long time.

The Housing Supply Statistics Information System (HIS) collects housing supply related data from the construction administration system and various housing information sites and calculates statistics. The system is built to quickly and precisely understand housing construction, sales results, and the state of unsold housing.

The Real Estate Announcement Prices (REAP) system provides the municipal entity with the announced price of the municipal house and collects the price of the individual house.

The system provides information such as individual single-family house, standard single-family house, apartment price, separate official land price, and official land price.

National Spatial Information Portal system (NSIP) is a system that integrates and manages the spatial information DB constructed by all public institutions. It is possible to analyze the data by linking the national standardized spatial information with the attribute information.

This paper proposes a method of linking the system through DB connection and API (Application Protocol Interface) or crawling public data. Also, this paper considers the direct linking through security protocol and data modification/loading method through ETL (Transformation, Loading).

2.3 Housing Market Analysis Technology

The housing market analysis methods can be divided into macro and micro analysis. In the case of macroeconomic analysis, we implement a VAR model that simultaneously analyzes multiple time series and a factor-augmented VAR (FAVAR) model that considered multiple variables at the same time in an ARIMA-type model that analyzes one variable in a time series. Also, reviewed the application of the DSGE (Dynamic Stochastic General Equilibrium) model. In addition to the macro analysis model, the micro analysis model is based on an economic base model, a shift-share model, a regional input-output model, and a computable general equilibrium (CGE) model.

In order to apply the above model to the platform, this system linked between the open source statistical software R and the commercial software E-Views for applying the existing model.

3 Direction of Construction of Open Platform Architecture

Open Platform is designed as a web-based open platform for collecting data from various sources and analyzing collected data easily and conveniently. In addition, the analytic model is stored on a component basis for scalability and reusability of the analytical model. In this way, this suggested system can collect, process, and analyze data based on the workflow form via the Web.

The data processing stage of the housing market integrated information platform consists of four stages of collection/storage/analysis/service infrastructure (Fig. 1).

The first step is to collect and input housing data, administrative data, spatial data, and other convergence data. In this step, ETL/Crawling technique is used to deal with the structured data.

In the second step, the data in the housing integration database is stored in Hadoop (Fig. 2). To be used for various model analysis such as time series analysis, micro and macro analysis. This method is suitable for time series analysis of large amounts of data.

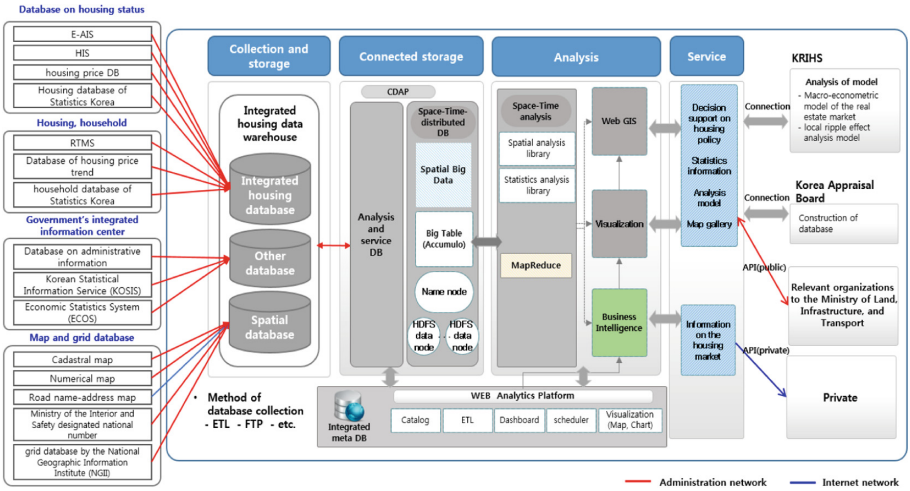


Fig. 1. Concept Diagram of Housing Market analysis platform.

In the third step, constructs a web-based analysis module and design a method to visualize the results. Users can analyze the housing market without any tools except web. In addition, by supporting user-based analysis components, frequently used models can be developed and utilized.

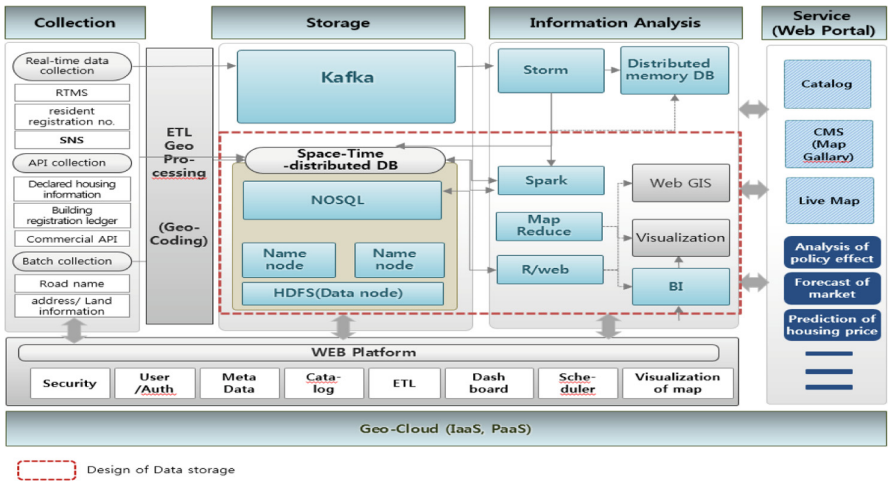


Fig. 2. Concept Diagram of Housing Market analysis platform

As a final step, the analyzed data is divided into user groups according to future policy directions such as private sector, researchers, and related organizations, and data and results are provided. In the case of public institutions, analytic components and

APIs are provided for statistical analysis and spatial analysis. In the case of private institutions, the analyzed results are provided in accordance with the needs of users to provide information.

4 NoSQL Based Space-Time Big Data Construction and Performance

In order to perform a large-capacity spatio-temporal big data analysis using an open platform, a technology capable of storing and processing data nationwide is needed. Therefore, this paper designed a NoSQL - based Spatiotemporal Big Data repository to store and process housing, price, and spatial information across the country.

The Spatiotemporal Big Data indexed the whole space of Korea based on Z-Curve to define a spatial index that can cover the whole space. This system designed and constructed grid-DB geo-coding through KEY-Value and a space-time repository to store information related to national housing (Fig. 3).

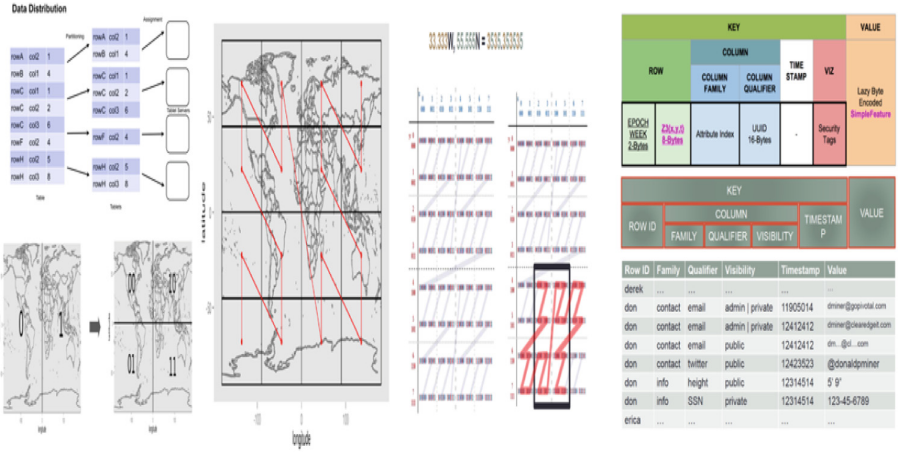


Fig. 3. Design of spatial index processing

The data block stored in Z-order curve according to the spatial index structure maximizes the search speed by having the search block located nearest to the spatial search. So, this paper measured the time of data storage in the Z-order curve method because saving time is important to store the data. The result confirmed the results better than the general database (Tables 1, 2).

Table 1. Loading performance test.

Test specification	Result of loading
– Servers: 4 – Core 24 physical, 48 Virtual – NIC: 1G – Disks: 6 1 TB – Memory: 256 GB	– Load 47 Table GBD (Geo Big Data) – Point 25, Line 18, Polygon 4 – 12Data type Point & Line, Point & Polygon, geometry-collection, etc.) – loading one Table (42 million data), time required is 11 min 19 s

Table 2. Detailed result of loading performance test.

Data description	Type	Number of data	Loading time	Remarks
Data of facilities cadastral map	Point	41,730,000	11 m~19 s	One table containing the maximum number
Hot spot of the official price of building	Point	360,000	11 m 44 s	Buffer generation
Buffer of telegraph pole	Point	9,200,000	5 m 4 s	Aggregate point data in 50 m buffer

5 Conclusion

In order to provide a reasonable policy decision support system for the change of the real estate market, it was necessary to construct a housing policy decision support system that can be easily and quickly accessed. This paper studied the implementation of open source platform, and designed the prototype NoSQL-based data repository. As a result, we confirmed the possibility of loading and processing large data in the housing market through open source programs. This paper verified the effectiveness of the spatial - temporal big data processing analysis of the structured data and to confirm the feasibility of the open platform for the analysis of the housing market. In the future study, we will propose a method to apply various models to support decision making through analysis of housing data and open platform.

Acknowledgments. The paper is supported by the "Korea Agency for Infrastructure Technology Advancement (KAIA)" project "Development of Housing Market Analysis and Forecasting Model Using Big Data (18RERP-B119172-03)".

References

1. Ok, J., Jo, M.: A study on big data analysis and utilization system of housing property (2015)
2. Lee, S.-H., Oh, J.: Development of integrated information platform for housing market analysis. In: Spring KIPS Conference, vol. 24, pp, 445–447 (2017)

Resource-Aware Migration Scheme for QoS in Cloud Datacenter

A-Young Son, DongYeong Son, Young-Rok Shin, and Eui-Nam Huh(✉)

Department of Science and Engineering,
Kyung Hee University, Yongin, Republic of Korea
{ayths28, ehddudl123, shinyr, johnhuh}@khu.ac.kr

Abstract. With the rapid growth of data centers, thousands of large data centers with lots of computing nodes are established. In order to user satisfaction, Assurance of QoS is important in CDCs. Also, many of the current research studies have not considered multi-metric for assurance of QoS. In this paper, we categorize QoS through previous work and build the migration scaling scheme for QoS in CDCs with considering multi-metric. And then from evaluation result, we prove that our proposed method is able to efficiently manage the resource and grantee QoS.

Keywords: Cloud datacenter · Resource management · Quality of Service

1 Introduction

For the popularity of the cloud data center, thousands of large data centers with lots of computing nodes are established. In order to meet their requirements and perform the services, it is necessary to assure QoS by service providers.

Researchers have developed mechanisms and systems to guarantee the QoS requirements of different services. However, many of the current research studies [1–4] have not considered multi-metric for QoS. In this paper, we propose an advanced migration scheme for QoS according to the current utilization of resources

Organization of the paper is as follows: Sect. 2 describes related work for QoS according to the objective in CDCs by analyzing previous approaches. Sect. 3 propose improved migration scheme for QoS. Sect. 4 provides migration efficiency evaluation and analysis. The last Chapter concludes this paper.

2 Related Work

2.1 Related Work for Energy Efficiency in CDCs

Table 1 summarizes resource management approaches for assurance of QoS. Researchers have developed mechanisms and systems to guarantee the QoS requirements of different services. In this paper, we intend to provide a deep insight into QoS issues by reviewing the technical details of QoS metrics and classifying them.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 85–90, 2020.
https://doi.org/10.1007/978-981-13-9341-9_15

Table 1. Related work for assurance of QoS.

QoS metric	Adaption method	Evaluation
Cost	Genetic algorithm	Aggregation function (sum, max, average)
Execution time, availability, price	Simple additive weighting	Aggregation function (sum, max, average)
Response time, cost	Linear programming	Aggregation function (sum, max, average)
Response time	Probabilistic slitting policy	M/M/1 Queuing model
Availability, average execution time	Queuing Model	Utility functions
Performance reliability	Markov model	Analytic solving of queue system

2.2 Categorization for QoS Metric According to Metric

Figure 1 shows a taxonomy of different approaches that are proposed in the literature to Assure of QoS in cloud datacenters. These QoS metric could be classified into three categories according to objectives: Cheapest Service, Best latency service, Best quality service. However, these approaches may lead to SLA violation [7, 8] and the additional cost of resource [2, 11].

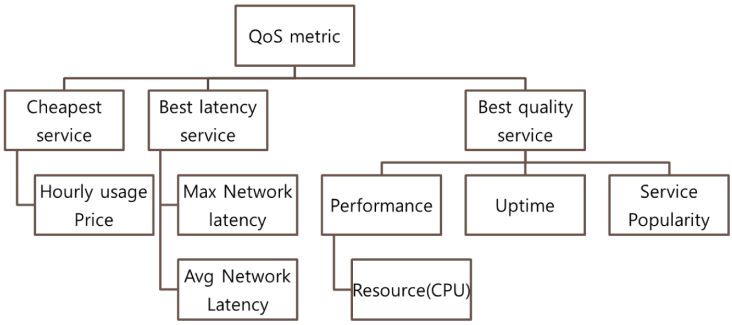


Fig. 1. QoS metric in CDC

3 Proposed Scheme for Assurance of QoS

3.1 Proposed Architecture

Figure 2 shows the detailed design of proposed scheme. It consists of three main modules: Service predictor Service Monitor and Migration manager.

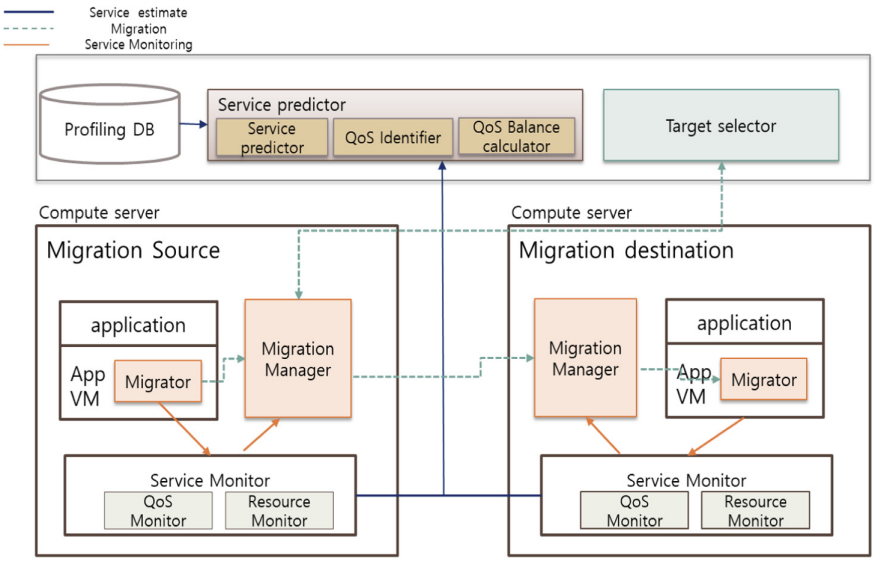


Fig. 2. Proposed scheme

1. Service predictor system: we determine VM status which needs to migration and selection of target machine according to service monitor information we applied to priority mechanism based on Eq. 1. In this module, we calculate a weight of QoS metric based rule. To find appropriate VM, The set of alternatives are ranked according to the Weigh value
2. Service monitor: It consists of two modules.
   A. Resource monitor: it is responsible for monitoring the resource status of both source machine and destination machines, including the resource utilization.
   B. QoS monitor: In this module, virtual machine configuration information which is essential to make migration decision based on resource monitoring information send to service predictor for prediction of service status
3. Migration manager: It is designed for making the decision on the VM migration situation whether or not. Depending on inputs such as QoS metric

3.2 Step of Proposed Scheme

Step of proposed scheme is as follows (Fig. 3):

Continuous resource monitoring is required in CDCs for resource management including the VM status information such as resource usage to make VM migration decisions

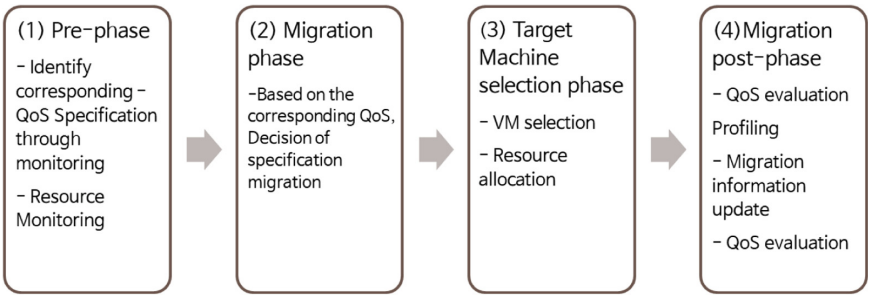


Fig. 3. Step of proposed scheme

Table 2. Notation in paper

Notation	Definition
$A_i(i = 1 \sim n)$	Alternative
$W_j(\sum w_j = 1)$	Weight
X_{ij}	j-th metric value of i-th alternative

Thus, for the selection of alternative (service), we define an alternative decision function that maximizes QoS using priority method as follows (Table 2)

$$A_i = \sum_j W_j X_{ij} \tag{1}$$

4 Evaluation

This chapter provides an evaluation of the proposed approach, and it also demonstrates how proposed scheme effectively. To compare Efficiency we use a simulation tool called Cloud sim 3.0. Our proposed schemes present advantages over the static schemes as was the number of deadlines missed smaller than those of static scheme like Fig. 4. Therefore, the proposed scheme proves the performance of migration by selecting VM according to the QoS metric.

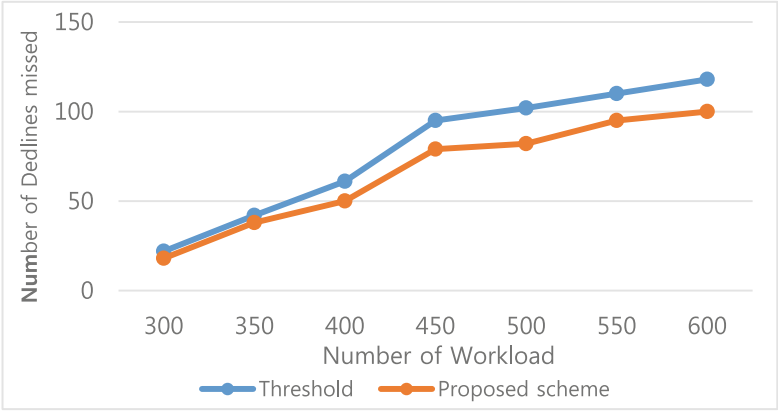


Fig. 4. Number of deadline missed for proposed scheme

5 Conclusion

Growing and expansion the of cloud service, we proposed a migration scheme for QoS. It is important to guarantee QoS for satisfying of user requirement. The goal of our advanced migration scheme is to assure QoS by determining migration and selection of VM.

According to the evaluation result, our proposed scheme is able to guarantee QoS and efficiently manage the resource in the cloud datacenter. The consideration of more general scaling models and other types of service to extend this work will be conducted in the future.

Acknowledgement. This research was supported by the MIST (Ministry of Science and ICT), Korea, under the National Program for Excellence in SW (2017-0-00093), supervised by the IITP (Institute for Information & communications Technology Promotion).

References

1. Beloglazov, A., et al.: A taxonomy and survey of energy-efficient data centers and cloud computing systems. Adv. Comput. **82**(2), 47–111 (2011)
2. Piraghaj, S.F., et al.: A survey and taxonomy of energy efficient resource management techniques in platform as a service cloud. In: Handbook of Research on End-to-End Cloud Computing Architecture Design, pp. 410–454 (2017)
3. Zhang, B., Sabhanatarajan, K., Gordon-Ross, A., George, A.: Real-time performance analysis of adaptive link rate. In: 33rd IEEE Conference on Local Computer Networks, 2008. LCN 2008, pp. 282–288. IEEE (2008)
4. Beloglazov, A., Abawajy, J., Buyya, R.: Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. Futur. Gener. Comput. Syst. **28**(5), 755–768 (2012)

5. Wu, C.-M., Chang, R.-S., Chan, H.-Y.: A green energy-efficient scheduling algorithm using the DVFS technique for cloud datacenters. Futur. Gener. Comput. Syst. **37**, 141–147 (2014)
6. Gunaratne, C., et al.: Reducing the energy consumption of Ethernet with adaptive link rate (ALR). IEEE Trans. Comput. **57**(4), 448–461 (2008)
7. Wu, G., et al.: Energy-efficient virtual machine placement in data centers by genetic algorithm. In: Neural Information Processing. Springer, Heidelberg (2012)
8. Maurya, K., Sinha, R.: Energy conscious dynamic provisioning of virtual machines using adaptive migration thresholds in cloud data center. Int. J. Comput. Sci. Mob. Comput. **2**(3), 74–82 (2013)
9. Graubner, P., Schmidt, M., Freisleben, B.: Energy-efficient virtual machine consolidation. IT Prof. **15**(2), 28–34 (2013)
10. Galloway, J.M., Smith, K.L., Vrbsky, S.S.: Power aware load balancing for cloud computing. In: Proceedings of the World Congress on Engineering and Computer Science, vol. 1 (2011)
11. Farooqi, A.M., Tabrez Nafis, Md., Usvub, K.: Comparative analysis of green cloud computing. Int. J. **8**(2) (2017)

A Dynamic FPGA Reconfiguration for Accelerating Machine Learning Framework with Image Service in OpenStack

Seungmin Lee(✉) and Sik Lee

Korea Institute of Science and Technology Information,
Daejeon 34141, Republic of Korea
{smlee76, siklee}@kisti.re.kr

Abstract. This paper discusses the feasibility of a dynamic FPGA reconfiguration by deploying FPGA images at run-time, with a particular focus on the offloading kernels to accelerate the training or inference phase of machine learning framework. It is useful to create FPGA images in advance and share them through repository because it takes a long time to build and is large in size. In this work, The Intel Arria10 GX FPGA card is used as a FPGA device, the TensorFlow as a machine learning framework and Glance, image service project in OpenStack, as a repository for storing and managing FPGA images. The OpenCL SDK tool chain presented by Intel corporation is used to implement FPGA images and client APIs for KeyStone and Glance are used to retrieve FPGA images from Glance. The result shows that our implementation can retrieve FPGA images, deploy them to the FPGA device and execute code that is offloaded to the FPGA device correctly.

Keywords: FPGA · OpenStack Glance · TensorFlow · OpenCL

1 Introduction

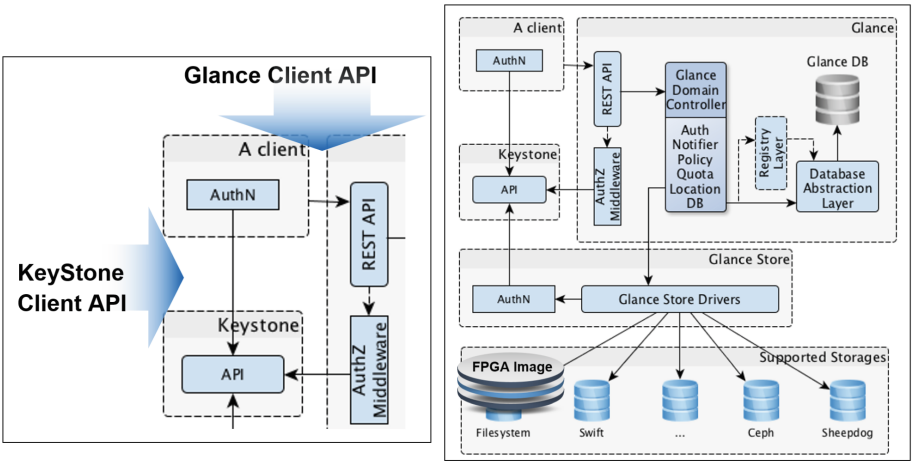
Artificial intelligence, deep learning, and big data analysis are one of the biggest issues, and there is a variety of open source frameworks to support them. In the aspect of computing resource, GPU is mainstream in model training, however, it has been increased to use field programmable gate arrays (FPGAs) in the field of machine learning because FPGA can achieve high performance with low-power [1, 2]. FPGA image, also known as partial reconfiguration (PR) bitstream, is an executable file and one should build and deploy it onto the FPGA device. In general, it takes a long time to build a FPGA image and is large in size. Therefore, it is useful to create FPGA images in advance and share them through repository. The Glance image service [6] in OpenStack provides storage not only for images of virtual machine but also for images of virtual disk, kernel and container. So, it is possible to extend the usage of Glance as a repository for FPGA images. This work examines the feasibility of a dynamic FPGA reconfiguration by deploying FPGA images at run-time, with a particular focus on the offloading kernels to accelerate the training or inference phase of machine learning framework. There are previous studies widely consider provisioning and virtualizing

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 91–96, 2020.
https://doi.org/10.1007/978-981-13-9341-9_16

FPGA when providing a regular virtual machine through OpenStack [3, 4]. The approach is similar to use Glance as a repository for FPGA image, however, our proposed solution can retrieve and deploy FPGA images so that dynamic FPGA reconfiguration is possible at run-time. This paper presents the following items that have been rarely reported by previous studies: (1) Implementation details of OpenCL host and kernel code for downloading FPGA images from Glance service, and (2) a method for calling functions of FPGA images at the TensorFlow.

2 Methodology

Figure 1 shows (a) a client for the OpenStack images API and (b) the architecture of Glance service [5]. Glance architecture consists of three parts: glance-api, glance-registry, and the storage for image. One can store and retrieve images with glance-api even though Glance doesn't supports FPGA image format because it doesn't check the contents of image format.



(a) Client API for KeyStone and Glance (b) OpenStack Glance Architecture

Fig. 1. Client for the OpenStack images API and architecture of Glance service.

The Glance API is a simple REST API for querying metadata or storing or retrieving images. It means that the metadata is returned as a JavaScript Object Notation (JSON) encoded mapping and image data is returned as binary format [6]. There is a Python API bindings for the OpenStack Image (Glance) API that can be used as a client for Glance service as shown in Fig. 1(a). One thing to note from Python API for Glance is that KeyStone client API is also required for authentication.

We should define the interface of a user-defined function to incorporate with the TensorFlow. It requires to specify types and names of input/output, and any attributes

the function might require. TensorFlow Python API provides *load_op_library* to load the shared library and register the operation with the TensorFlow framework.

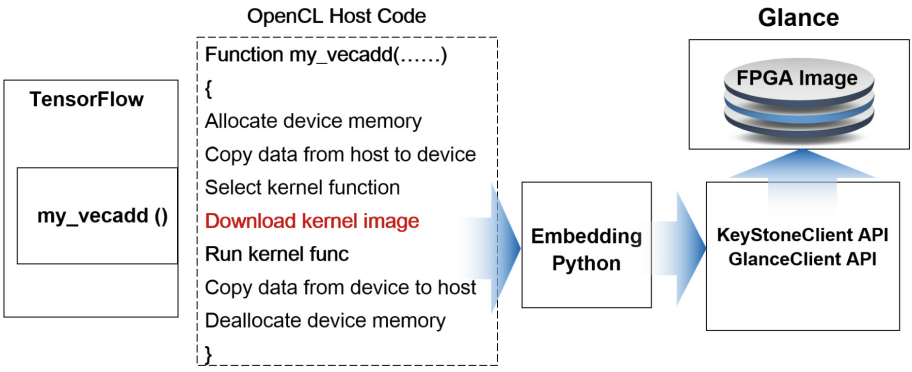


Fig. 2. Flow of FPGA image retrieving from Glance and combining to the TensorFlow.

Figure 2 shows the complete flow of operations performed by our proposed system to take FPGA image and deploy it on a FPGA. The very first step in the flow is to take the OpenCL host code and retrieve kernel image by calling functions in the Embedding Python module. The Embedding Python module is required because the client API for KeyStone and Glance exposes Python interface whereas the implementation of user-defined functions to the TensorFlow should be written in C++ with following some prerequisites and rules [7]. Therefore, the Embedding Python module should exist among TensorFlow, OpenCL host code, and client API for KeyStone/Glance though the OpenCL supports both C/C++ and Python.

```
kernel void vecadd(global float *a, global float *b,
global float *c)
{
    int tid = get_global_id(0);
    c[tid] = a[tid] + b[tid];
}
```

List. 1. Sample Kernel Code included in the FPGA Image

List. 1 shows the OpenCL kernel code, where we offload the vector addition section to FPGA. The kernel is implemented using the OpenCL SDK tool chain presented by Intel corporation. The tid is the index of OpenCL thread, and element-wise addition of the two input vectors is performed.

3 Implementation and Results

We implement modules as mentioned in the previous chapter and generate FPGA image from List. 1 kernel code by compiling Intel FPGA SDK for OpenCL version 18.0 Prime Pro. The test system is set up with 2.2 GHz Intel Xeon Silver 4114 processor, 64 GB DDR4 memory, and a PCIe-based FPGA card, Intel Arria10 GX FPGA. CentOS 7.5.1804 with Linux kernel 3.10.0-862 is chosen as the operating system.

```
void my_vecadd(float *A, float *B, float *C)
{
  // retrieving FPGA image
  loadBinaryFromGlance(fpga_img, &img_size);
  program = clCreateProgramWithBinary(context, 1, &device,
      &img_size,(const unsigned char **) &fpga_img, 0, 0);
  clBuildProgram(program, 0, NULL, "", NULL, NULL);
  clEnqueueNDRangeKernel(queue,kernel,1,0,&gs,&ls,0,0,0);
}
```

List. 2. The code snippet of OpenCL host program for retrieving FPGA image and deploying to the FPGA device

The code snippet shown in List. 2 describes the way to retrieve the FPGA image using *loadBinaryFromGlance* and deploy it to the FPGA device with OpenCL APIs, *clEnqueueNDRangeKernel*. *clCreateProgramWithBinary* enables to reuse the output of compilation, thereby we can execute the FPGA image retrieving from image repository. We note that *clBuildProgram* must be called for program created using *clCreateProgramWithBinary* to build the program executable for FPGA device [8].

```
void loadBinaryFromGlance(char *buf_img,size_t *buf_size)
{
   pModule = PyImport_ImportModule("glanceclient2fpga");
   pGet_aocx = PyObject_GetAttrString(pModule,"get_aocx");
   pRetValue = PyObject_CallObject(pGet_aocx, pArgs);
   // get buffer data
   while((item=PyIter_Next(iterator)) != NULL) {
      size_t item_size = PyString_Size(item);
      char *buf = PyString_AsString(item);
      memcpy(&buf_img[recv_size], buf, item_size);
      recv_size += item_size;
   }
}
```

List. 3. The code snippet of Embedding Python Module to convert data from python structure to C/C++ buffer

List. 3 shows the code snippet of Embedding Python Module. As there are many ways accessing Python from C, we choose to get a reference to the main module and namespace by importing Python library modules with *PyImport_ImportModule*, and

then calling Python functions, classes, and methods with *PyObject_CallObject*. The code in List. 3 has a while loop because the Python data structure containing FPGA image is an iterable object which means we can access data in the iterable object with next method, in this case *PyIter_Next*.

```
# glanceclient2fpga.py file
from keystoneauth1 import loading
from keystoneauth1 import session
from glanceclient import Client

def get_aocx(auth_url,user_id,passwd,prj_id,img_id):
    loader = loading.get_plugin_loader('password')
    auth =loader.load_from_options(connection_information)
    session = session.Session(auth=auth)
    glance = Client('2', session=session)
    img_data = glance.images.data(img_id)
       return (img_data.iterable, img_data.length)
```

List. 4. Read Intel FPGA Executable (aocx) image using KeyStone and Glance client API

List. 4 shows the code snippet about receiving image data using KeyStone and Glance client APIs. The extension of executable file for Intel FPGA is aocx and Glance service allows any file extension, thereby we can retrieve FPGA image with this extension. The *get_aocx* function returns iterable object for image data with size of it as mentioned in the List 3.

We use the same hardware for two machines running OpenStack KeyStone and Glance services, and the machine learning framework, TensorFlow. Each is set up with 2.2 GHz Intel Xeon Silver 4114 processor, 64 GB DDR4 memory, in particular a PCIe-based FPGA card, Intel Arria10 GX FPGA is equipped for the TensorFlow. CentOS 7.5.1804 with kernel 3.10.0-862 is chosen as the operating system. The GNU C compiler gcc 4.8.3 is used to build TensorFlow 1.9.0 from source code and Python 2.7.15 is used to execute the test code shown in List 5.

```
import tensorflow as tf
my_module = tf.load_op_library("./my_vecadd.so")

v1=tf.constant([i for i in range(1,101)],dtype=tf.float32)
v2=tf.constant([i for i in range(1,101)],dtype=tf.float32)
output = my_module.my_vecadd(v1, v2)

sess = tf.Session()
print(sess.run(output))
```

List. 5. A test code to verify the correct execution of a dynamic FPGA reconfiguration

The List. 5 shows the TensorFlow test code as a proof of concept. We repeat the test cases by changing the value of vectors, *v1* and *v2*, to verify the correct execution of our FPGA reconfiguration at run-time. Finally, we can get the same results running at the host CPU throughout all test cases.

4 Conclusion

This work explores the feasibility of a dynamic FPGA reconfiguration by deploying FPGA images at run-time, with a particular focus on the offloading kernels to accelerate the training or inference phase of machine learning framework, TensorFlow. The PCIe-based Intel Arria10 GX FPGA card with the OpenCL SDK tool chain is used to verify correctness of execution and to generate FPGA image, respectively. We also show that Glance image service in OpenStack can be extended as a repository for FPGA images.

References

1. Suda, N., Chandra, V., Dasika, G., Mohanty, A., Ma, Y., Vrudhula, S.B.K., Seo, J., Cao, Y.: Throughput-optimized OpenCL-based FPGA accelerator for large-scale convolutional neural networks. In: Proceedings of the ACM/SIGDA ISFPGA, pp. 16–25 (2016)
2. Utku, A., Shane, O., Davor, C., Andrew, C.L., Gordon, R.C.: An OpenCL deep learning accelerator on arria 10. In: International Symposium on Field-Programmable Gate Arrays, 2017, Monterey, California, USA (2017)
3. Byma, S., Steffan, J., Bannazadeh, H., Garcia, A.L., Chow, P.: FPGAs in the cloud: Booting virtualized hardware accelerators with OpenStack. In: Proceedings of the IEEE 22nd International Symposium on FFCCM, pp. 109–116, May 2014
4. Chen, F., et al.: Enabling FPGAs in the cloud. In: Proceedings of the 11th ACM Conference on Computing Frontiers, pp. 3:1–3:10 (2014)
5. OpenStack Glance Architecture. https://docs.openstack.org/glance/pike/contributor/architecture.html
6. Pepple, K.: Deploying OpenStack. O'Reilly Media, Sebastopol (2011)
7. Adding New Operations to TensorFlow. https://www.tensorflow.org/extend/Adding_an_op
8. Khronos OpenCL Working Group, The OpenCL Specification, version 1.2.19, November 2012. https://www.khronos.org/registry/cl/specs/opencl-1.2.pdf

Switchless Interconnect Network with PCIe Non-Transparent Bridge Interface

Sang-Gyum Kim[1], Yang-Woo Lee[1], Seung-Ho Lim[1(✉)], and Kwang-ho Cha[2]

[1] Division of Computer and Electronic Systems Engineering, Hankuk University of Foreign Studies, Seoul, South Korea
lim.seungho@gmail.com
[2] Center for Supercomputer Development, Korea Institute of Science and Technology Information, Daejon, Korea

Abstract. HPC systems are one of the important role to run a lot of high-performance applications. Conventionally, HPC deploys switch-based interconnect networks to connect host nodes with high cost and high performance technologies such as Infiniband and Ethernet. However, high cost interconnection switches may not be required by desired cost-effective HPC system. To give cost-effective interconnect network interface, we introduce prototype of switchless interconnect network with PCIe NTB to give cost-effective interconnect network for HPC systems. In the prototyped system, computing nodes are connected to each other via PCIe NTB interface constructing switchless interconnect network such as ring or 2d mesh network. Also, we have implemented prototyped data sharing mechanism on the switchless interconnect network system. The designed switchless interconnect network system is cost-effective as well as it provides competitive data transferring bandwidth within the interconnect network.

Keywords: HPC · Interconnect network · Non-Transparent Bridge · PCIe · Switchless

1 Introduction

Nowadays, HPC systems are one of the important role to run a lot of high-performance applications such as big data and cloud computing-based applications. In the high performance computing (HPC) systems, hundreds of computing nodes are connected to high performance interconnection network to get high level performance such as high throughput and low latency. As computing and network technology is developed rapidly, the amount of data for processing and transferring between computing nodes in HPC is rapidly increased, which results in requirements of high performance. Meanwhile, cost and power consumption is also important point for developing HPC systems with the rapid increase of number of connected computing nodes and memory, and peripheral I/O devices such as storage.

The most HPC systems deploy interconnection network with Infiniband, Ethernet, and PCI Express based devices [1]. Although, the traditional interconnect networks can do

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 97–102, 2020.
https://doi.org/10.1007/978-981-13-9341-9_17

high bandwidth, their cost and power dissipation is getting larger with the increase of volume of HPC. Most of the interconnect network connects nodes via interconnect switch and adapter cards between nodes. The switch-based architecture does support good network topology, it gives high cost for developing interconnect network and HPC systems.

On the other hand, PCI Express technology was originally developed to support of connection among processors and peripheral I/O devices in a computing system [2, 3]. It has been developed. PCI connects processor and peripheral devices with a transparent bridge manner for processor manages addresses of peripheral devices as a master. While PCI bridge connected as a transparent fashion between processor and devices. PCI Non-Transparent Bridge (NTB) is a function that translates PCI transactions from one PCI hierarchy into corresponding transactions in another PCI hierarchy. During for years, PCI NTB acted as a connection between two PCI-based systems by mapping some address region to another each other [4, 5]. Namely, PCI NTB do support high cost-efficient multi-host system between two host with only single non-transparent bridging interface.

In this paper, we introduce prototype of switchless interconnection network using PCIe NTB to give cost-effective interconnect network for HPC systems. In the prototyped system, computing nodes are connected to each other via PCI NTB interface constructing switchless interconnect network such as ring or 2d mesh network. We implemented data sharing mechanism within the PCI NTB-bases switchless interconnect network, and showed its feasibility and possibility as a cost-effective and high performance interconnect network.

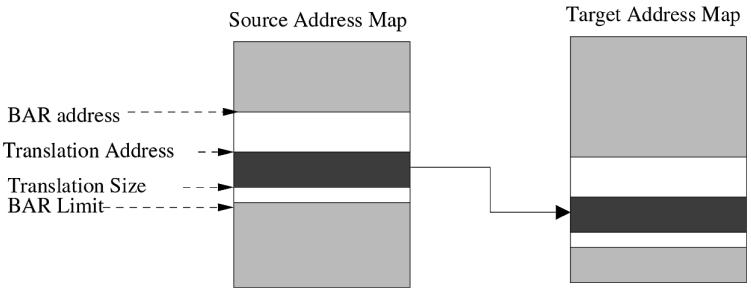


Fig. 1. PCIe NTB address translation mechanism

2 PCIe NTB and Switchless Interconnect Network

PCI non transparent bridge is functionally similar to transparent bridge except that there is a process on both sides of the bridge, in which each processor has its own independent address space domain. In the NTB environment, PCIe bridge translate addresses that across from one memory address space to other space through translation registers, which is so called memory window. With this address translation two processors connecting through PCI NTB can share independent address space, each of which is belonging to each processor, and transfer data through this independent address space. The address translation of one host to another host is described in Fig. 1.

In addition to that, two processors can share some small data set through some special registers, i.e., eight or more 32 bit registers, called ScratchPad registers. When one host write data to one of ScratchPad registers, other host can access and read directly from the ScratchPad register. Also one processor can cause an interrupt to the other processor through register called Doorbell register. There are sixteen doorbell interrupts that can be set or cleared, as well as masked.

Basically, NTB provides address translation between two separate address space from two host connected. To make up switchless interconnect network for multiple hosts system, in this paper, two or more NTB host adapter is connected to a single-CPU-based host, and the NTB host adapter is connected to each other. For instance, the ring topology interconnected multi-host system is described in the Fig. 2. As shown in the figure, each host can share separated address space trough NTB address translation on both sides of the host connected. At the initial state, in each host, a host Id is assigned and address space is allocated to be used and shared in the network. Then, the information for address region allocated, that is address offset and size, is transferred to its neighbors. With those information, each host can set the address regions of neighbors that are connected to itself, which is used for data sharing and transferring.

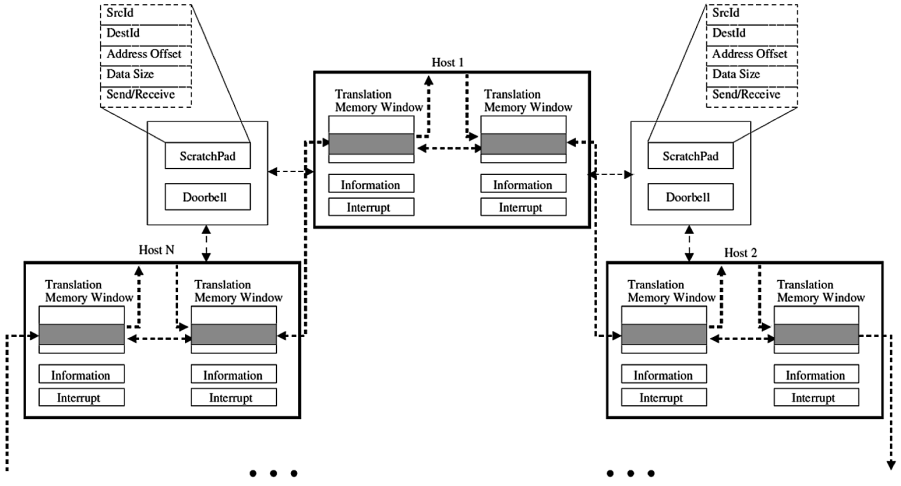


Fig. 2. Switchless interconnect networking and data sharing with PCIe NTB

After the initial state, each host can transfer data through shared address space, as a result, can send or receive data to destination host or from source host with specified host Id. The data transfer is done as follows. When a host tries to send data to other host, it writes data from its own memory space to within shared address space region, Data can be written DMA [7, 8] or directly memcpy. When data is being written, destination address is translated to address region of the connected side via NTB translation. After data is written, a host send information such as SrcId, DestId, Address offset, Data size, and flag for Send/Receive through ScratchPad registers. Then a host

triggers an interrupt with doorbell register that is alert for transferring data. When connected neighbor host receive an interrupt triggered by a sending host, it identifies that who is source host and who is destination by reading ScratchPad registers. If the destination host is himself, the host gets data directly from the shared address region, and transferring is done. If the destination host is not him, the host should bypass the received data to its another connect host. The bypassing mechanism is similar to described above. Those transferring sequence is done when destination host gets data from its connected host. By those operations, data can be shared multiple hosts connected NTB-based switchless interconnect network.

3 Evaluation

The prototype platform of switchless interconnect network based on PCIe NTB is setup with three processing nodes having Intel Core-i7 3.7 GHz CPU and 8 GB DRAM. For PCIe NTB connection, PLX technology's PEX 8749 and 8733 Chipset-based prototyped PCIe host adapter is used [10]. For each host, PEX Chipset-based adapter is connected and those are connected each other to make ring networks. In host system, Linux Operating System with kernel version 4.16.3 is run with the NTB PCIe device driver, that is PEX 8x NTB device driver, which supports PCIe Gen3 specification and DMA support. The NTB adapter supports up to four, eight, and sixteen PCI channels to provide data transferring bandwidth. We have setup ring network with three hosts, in which each host installed two NTB host adapters. The NTB host adapter is connected to each other to make ring network with PCIe fabric cable that supports PCIe Gen3 eight channel. Based on this prototyped environment, we have implemented data sharing mechanism among processing nodes connected to the switchless networks.

For the experiments, we evaluated data transfer rate using DMA data transfer between nodes within the switchless network. The Fig. 3 represents experimental results for DMA data transfer in PCIe NTB-based ring network, while it is compared with single DMA connection. The NTB device, used in this experiment, provides up to ranging from 20Gbps to 30Gbps between two independent host system through NTB data transferring according to the PEX chipset and connection environment, as shown in the Fig. 3(a). On the contrary, Fig. 3(b) represents data transfer rate for each host interface when all the hosts communicate simultaneously. From Fig. 3(b) we identify that data transfer rate is slightly diminished by the simultaneous transfer in the ring in comparison with independent data transfer between two hosts. It implies that overall data transferring rate can increase as the number of nodes increase since overall data sharing rate can be said to it is total amount per node. The Fig. 3(c) shows how much data rate is diminished due to the simultaneous transferring in the ring network in comparison with single connection. As shown in the figure, the overhead is small, while overall network throughput in the ring network is increased as hosts nodes increase. In summary, the prototyped system shows that cost-effective switchless interconnect networked cluster with PCIe NTB can share data with relative higher data transfer rate.

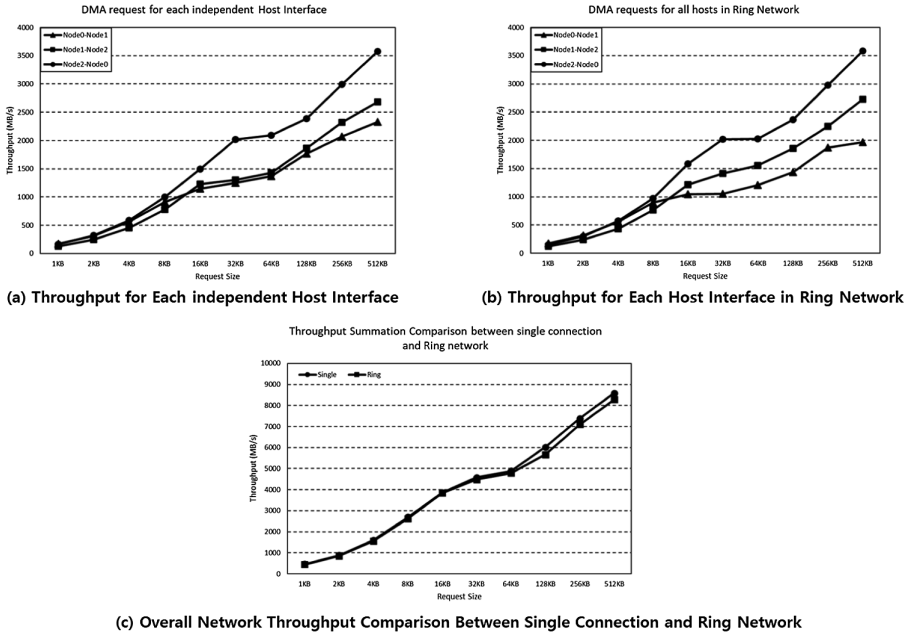


Fig. 3. Experimental results for DMA data transfer in NTB-based ring network in comparison with single NTB interface.

4 Conclusion

In the High Performance Computing (HPC) system, interconnect network is one of the key components since data interchange and sharing is one of the important parts for HPC. Conventionally, HPC deploys switch-based interconnect networks for connecting host nodes using high cost and high performance technologies such as Infiniband and Ethernet. However, high cost interconnection switch may not be required by desired cost-effective HPC system. In this paper, we designed and implemented switchless interconnect network prototype. In the designed interconnection network, PCIe Non-Transparent Bridge (NTB) technology are used to connect computing nodes. NTB is PCIe bridging technology that can provide separate address region connecting to the two nodes. We setup prototyped switchless interconnection networked system using multiple hosts connecting with NTB among nodes. Also, we have implemented prototyped data sharing mechanism on the switchless interconnect network system. The designed switchless interconnect network system is cost-effective while it provides competitive data transferring bandwidth within the interconnect network.

Acknowledgement. This research also has been performed as a subproject of Project No. K-18-SG-27-03R-1(Development of key enabling technologies for massively parallel and high-density computing) supported by the KOREA INSTITUTE of SCIENCE and TECHNOLOGY INFORMATION (KISTI), 2018. The Research was supported by the National Research

Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (NRF-2016R1C1B2009914). This work was supported by Hankuk University of Foreign Studies Research Fund.

References

1. Top500.org: Interconnect Family Statistics [Internet] (2015). http://top500.org/statistics/list
2. Helal, A.A., Kim, Y.W., Ren, Y., Choi, W.H.: Design and implementation of an alternate system inter-connect based on PCI Express. J. Inst. Electron. Inf. Eng. **52**(8), 74–85 (2015)
3. Liu, J., Mamidala, A., Vishnu, A., Panda, D.K.: Evaluating infiniband performance with PCI Express. IEEE Micro **24**(1), 20–29 (2005)
4. Heymian, W.: PCI Express multi-root switch reconfiguration during system operation. M. Eng. thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge (2011)
5. Krishnan, V.: Towards an integrated IO and clustering solution using PCI express. In: 2007 IEEE International Conference on Cluster Computing, pp. 259–266 (2007). http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4629239
6. Mohrmann, L., Tongen, J., Friedman, M., Wetzel, M.: Creating multicomputer test systems using PCI and PCI Express. In: IEEE AUTOTESTCON, pp. 7–10 (2009). http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=5314043
7. Choi, M., Park, J.H.: Feasibility and performance analysis of RDMA transfer through PCI Express. J. Inf. Process. Syst. **13**(1), 95–103 (2017)
8. Rota, L., Caselle, M., Chilingaryan, S., Kopmann, A., Weber, M.: A new DMA PCIe architecture for Gigabyte data transmission. In: 2014 19th IEEE-NPSS Real Time Conference (RT), pp. 1–2 (2014). http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7097561
9. Richter, A., Herber, C., Wild, T., Herkersdorf, A.: Resolving performance interference in SR-IOV setups with PCIe Quality-of-Service extensions. In: Euromicro Conference on Digital System Design (2016)
10. ExpressLane PEX8749 PCI ExpressGen 3 Multi-Root Switch with DMA Data Book, PLX Technology (2013)
11. Shim, C., Cha, K., Choi, M.: Design and implementation of initial OpenSHMEM on PCIe NTB based cloud computing, Cluster Comput. (2018). https://doi.org/10.1007/s10586-018-1707-0
12. Respondek, J.: Numerical approach to the non-linear diofantic equations with applications to the controllability of infinite dimensional dynamical systems. Int. J. Control **78**(13), 1017–1030 (2007)

A Proposal of IoT Message-Oriented Service Framework for Serverless Software Architecture

Sunggeun Yoo, Minjeong Song, and Sangil Park(✉)

Department of Information Technology and Media Engineering,
The Graduate School of Nano IT Design Fusion, Seoul National University
of Science and Technology, Seoul, Republic of Korea
orcogre@gmail.com, mi0_0ng@naver.com,
{gmpark,sipark}@seoultech.ac.kr

Abstract. The serverless software architecture is a hot topic in software architecture world. The architecture is suitable for many applications, such as data validation and real-time data analysis for Internet of Things (IoT). Nowadays, many applications are encountered special situations: IoT applications will bring many events and processes will be prepared for lots of events. We designed the new architecture which uses Internal Message Passing Application Protocol Interfaces (IMPAPIs). The interface adopted Advanced Message Queueing Protocol (AMQP). The design reduces assigned processes which have waited for events. Our suggested architecture is suitable for event-driven service situation.

Keywords: Serverless · AMQP · Cloud computing · Message queue · Event-driven

1 Introduction

With the proposal of cloud computing and Service Oriented Architecture (SOA), a new system structure called Function as a Service (FaaS) emerged. This structure is the basis of an architecture called serverless. When an event occurs, the backend system generates a process to process the operation asynchronously with the function. This idea has several advantages. FaaS has an advantage that server system management or server application is not required [1].

Also, these processes rarely use shared memory compared to existing service architectures. This structure has the advantage that a simple and robust system can be constructed. And, in case of performing a task requiring a lot of computing resources, it is necessary to control the order of each function. In the case of a location-based tracking system, the sequence of events is important [2].

For this sequence control, the Service Harmonizer as shown in Fig. 1, harmonizes all event sequences. Also, the service harmonizer collects the response result of the function returned to the callback and transfers it to another function to produce a meaningful result.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 103–107, 2020.
https://doi.org/10.1007/978-981-13-9341-9_18

In this paper, we propose a new architecture suitable for Open Complex Service Framework and show that this architecture is suitable for IoT or other cloud-based service applications.

2 Proposed Service Framework

The classification of the services to be provided to implement the proposed service framework is very important. In general, services provided by the cloud can be classified into two types: stateful services and stateless services.

A stateless service is a service whose processing time is short, such as providing a user authentication or a web page, and a state in which the processed state can be instantly expressed. Also, the processing state of such a service is managed on the client side. The stateless service can be directly connected to the Representational State Transfer Application Programming Interface (REST API) gateway to provide immediate response.

On the other hand, stateful services often have complex processing structures. If a user requests a map service or a push service to track his/her movement in real time, an event generated by the user is inserted into the server in real time, and the server must return a response in real time.

A system created with a common server-client structure should control the response by creating a thread to control each event and its processing. However, the proposed framework can do this without creating a thread. have. In other words, the proposed system puts a sequence of events that need to be controlled by the service harmonizer in the queue and sends the result back to asynchronous callbacks in order and sends the result to another function that gets the result to get the whole result.

Such a design can shows many possibilities. In particular, it has a great advantage in implementing a cloud-based editing program that requires a long processing time, such as image processing and audio processing, and processes complicated calculations

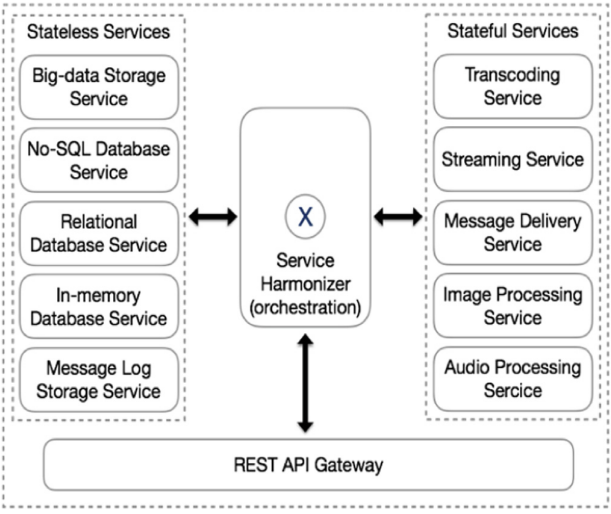


Fig. 1. The whole configurations of proposed service framework.

according to user input. In addition, the proposed system uses Advanced Message Queuing Protocol (AMQP) as a message protocol for interface between internal functions. The overall framework is shown in Fig. 1.

3 Proposed Message Formats

In a server system based on an asynchronous processing such as an existing cloud, a process occurring due to a sequential event has to be processed using a complex concurrency control logic. The present invention proposes a message format for transmitting and processing an instruction in a message in an asynchronous based server. Basically, messages are structured information such as XML (eXtensible Mark-up Language) or JavaScript Object Notation (JSON) with a hierarchical structure. These messages are vulnerable to errors, and errors are likely to propagate throughout the message structure. Therefore, the proposed message structure is shown in Fig. 2.

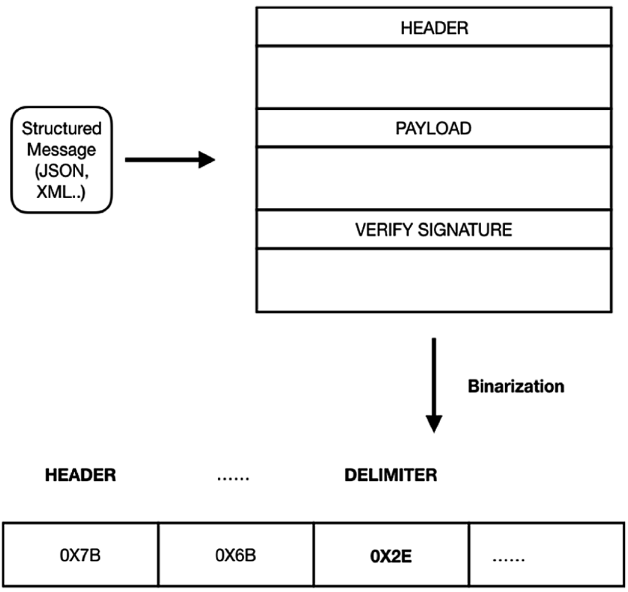


Fig. 2. Diagram of proposed message format

A message representing an event is divided into Header, Body, and Verify Signatures, such as a network packet, and Verify Signature has a Summery that detects if a message has failed. Header stores the sequence of events as a number as shown in Fig. 3. It also has a system as shown in Fig. 4 for handling this message. The structure of proposed message format shown in Fig. 5.

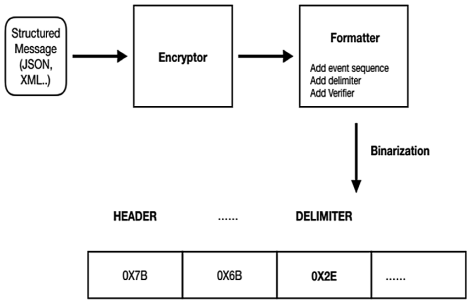


Fig. 3. Encoding method of proposed message format.

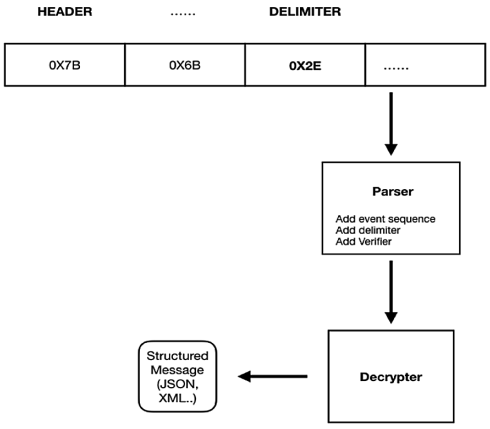


Fig. 4. Decoding method of proposed message format.

Fig. 5. Structure of proposed message format.

4 Conclusion

The framework proposed in this paper has a structure that is well suited for event-based IoT applications and cloud-based video editing applications. In particular, if the size of the system is large and large-capacity processing is required, the effort and cost for constructing the server system can be greatly saved. To test the proposed framework, we are building a test bench using the RabbitMQ message framework that implements AMQP and the cloud platform based on Infrastructure as a Service (IaaS). We expected the framework that will be used as a render system for dynamic 3D object reconstruction or cognitive IoT systems.

Acknowledgement. This work was supported by 2018 Project to Support Leading Model for Commercial-ization with a University (Implementation of the Academic-Educational Cooperation System based on the Academic Linkage through the Establishment of Hongneung Tech-Biz OPERA Center).

References

1. Roberts, M.: Serverless Architectures. http://martinfowler.com/articles/serverless.html
2. Yoo, S., Park, S.: A proposed message format for serverless IoT software architecture. In: Proceedings of the 4th International Symposium on Advanced and Applied Convergence, November 2016

Deep Learning Based Gesture Recognition System for Immersive Broadcasting Production

Meeree Park, Sung Geun Yoo, Minjeong Song, and Sangil Park(✉)

Nano IT Design Fusion Graduate School, Seoultech, 232 Gongneungno, Nowon-Gu, Seoul 01811, Republic of Korea
{alfl8514, mi0_0ng}@naver.com, orcogre@gmail.com, sipark@seoultech.ac.kr

Abstract. In this paper, we implement a system that provides the convenience of personal broadcasting production by recognizing the user's operation using the sensor tag and implementing the function corresponding to the operation. The system can acquire sensor data and learn the data through deep learning to distinguish the user's gesture. In this paper, we study the process of recognition of data through the data acquisition process and the deep learning process using the sensor tag and propose a method to perform the function using it.

Keywords: Gesture recognition · Deep learning · Immersive broadcasting · MEMS sensor · Internet of things

1 Introduction

Recently, a variety of immersive experience applications such as VR and AR are emerging due to development of immersive media technology. However, the development of immersive experience applications has the disadvantage that many people and equipment must be mobilized. To solve this problem, various tools for improving the convenience of media production have been developed. Among them, gesture recognition technology is being explored in order to improve accuracy by using 2D image due to development of technology such as deep learning. However, there are disadvantages of using images that are less accurate, such as failure to detect a person's hand or untrained images, because only simple image data is treated with information [1]. Due to the recent miniaturization of wearable devices, a variety of devices such as Galaxy Gear, Apple Watch and Xiaomi Mi Band have emerged, making it easy to acquire a variety of data. In particular, sensor data such as acceleration and gyro can be used to obtain direction and angle of movement. This information is transmitted through Bluetooth and ZigBee, and when the gesture recognition is performed by using the deep learning framework, information according to the gesture can be transmitted and a specific function can be performed. Therefore, in this paper, we have studied a gesture recognition framework that can be used for realistic broadcast production using gesture recognition and sensor network.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 108–115, 2020.
https://doi.org/10.1007/978-981-13-9341-9_19

2 Related Works

There has been a lot of interest in the area of IoT sensor networks with wearables in the recent year [2–5]. Sensors used to IoT has been developed smaller, so many industries applicate to various that including mobile devices, clothes, fitness trackers [6–8].

We were inspired by the recognition of the situation through the recognition of hand gestures. There are a variety of ways to analyze gesture recognition data now. One of these representative methods is based on Vision [9, 10]. However, vision-based hand gesture recognition is not only troublesome to install cameras and lights but also has a disadvantage that it is vulnerable to the position and lighting condition of the camera. In our study, hand gesture recognition is performed using a wearable device based on a sensor tag, so that broadcasting can be freely performed without being dependent on an external environment.

As a method of analyzing the result data of the sensor, the logistic regression using the predictive model is widely used through the independent variable and the dependent variable [11]. In recent years, research has been actively conducted through data analysis using machine learning rather than regression analysis. SVM (Support Vector Machine), which is one of the machine learning techniques, is now the most widely used [12]. In this paper, we use deep learning as a method of analyzing sensor tag data. By using this deep learning, we can achieve higher accuracy than using conventional logistic regression or SVM, and we expect that real-time data processing will be possible through a web-based deep learning framework.

3 Proposed Method

We constructed an experimental framework to distinguish hand gestures with gyro sensor and acceleration sensor data among the sensor data acquired with sensor tags attached to both hands. Yang et al. [3] developed a three-axis accelerometer that acquires data from a smartphone and performs a study to distinguish between dynamic postures and static postures, sitting, standing, walking, and running [Real-Time Physical Activity Recognition Using Tri-axis Accelerometer of Smart Phone].

For sensor data acquisition, two sensortag were attached to both hands as shown in Fig. 1. The use of the two is to attach one more backup in case the communication is disconnected, and the actual behavior classification accuracy is based on only one of the data.

The scenario of the gesture recognition system implemented in this paper is as follows. Using the Sensor Tag as a wearable device, wear it on the user's wrist and perform a specific gesture. Through this process, the gyro and the acceleration value can be obtained. The gesture is labeled by measuring the start time and end time of the gesture through the image data of the webcam that synchronizes the data with the time.

The sensor data is acquired at 10 sampling rates per second, and is transmitted to the router through the Bluetooth BLE, and the router is transmitted to the Amazon Lamda cloud through the LTE Modem. A framework set for analyzing the data was constructed by storing the acquired data in a file type database, sqlite. As shown in Fig. 3, the recorded video timestamp is synchronized with the time of the router, so that

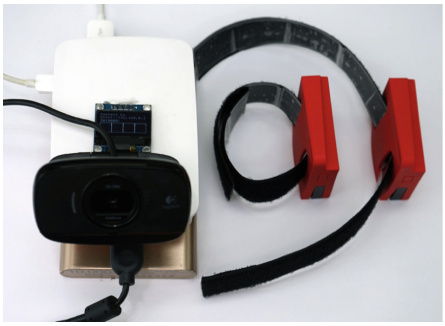


Fig. 1. Bluetooth BLE Router for acquiring sensor data (left), and two sensortag with detachable wristbands (right)

the sensor data obtained and the behavior of the image are exactly matched. As shown in Fig. 4, the gesture of the operator is defined, and labeling is performed in accordance with the recorded gesture. The recorded acceleration and gyro sensor data are synchronized with the timestamp, The gesture of the time that was there was read and classified by the deep running framework (Fig. 2).

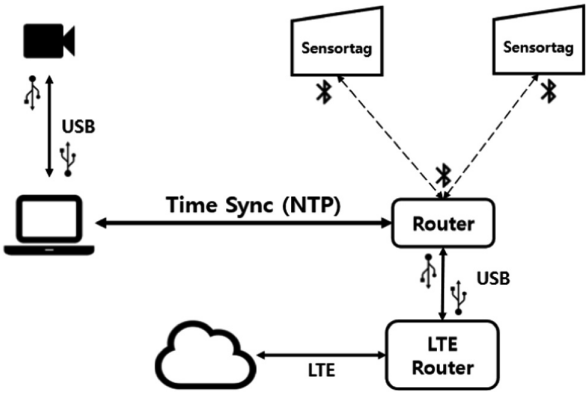


Fig. 2. The whole hand gesture analysis framework configuration for acquiring sensor data.

The acceleration and gyro data structures obtained from the sensor tags are shown in Fig. 3. Acceleration When accMeter table and event table are joined and the time that data entered in accMeter table is between the start time and end time of the event, we assumed it is the state of the event. The gyro data classifies the state of the gesture in the same way.

The gesture defined in this paper is shown in Fig. 4. For data collection, the Sensor tag CC2650 from Taxas Instruments was worn on both arms. One experimenter performed 100 experiments for each of five motions. This information can be learned by using deep learning to recognize a user's action and execute a function corresponding

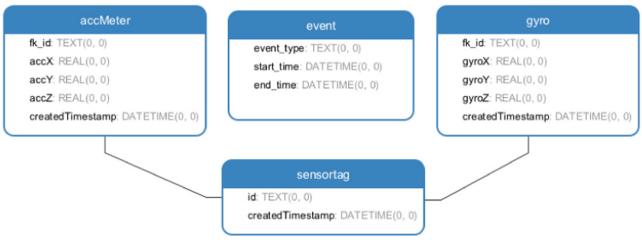


Fig. 3. Data structure of sensor tag

to the learned action. The five actions specified in this paper support convenience of production by starting broadcasting, ending broadcasting, sound effects, pausing of broadcasting, and switching cameras.

Icon	Gesture	function
	Lift the right arm.	Start broadcasting.
	Shake the right hand.	Broadcast end.
	Clapping.	Sound effect.
	Cross the arm.	Pause.
	Swing the right hand.	Switch camera.

Fig. 4. One kernel at x_s (*dotted kernel*) or two kernels at x_i and x_j (*left and right*)

4 Experimental Results

In this experiment, three-axis acceleration and gyro sensor data sent from a sensor tag attached with a bracelet to both hands are sent to a deep learning framework to learn the gesture of a broadcasting producer who performs a predetermined gesture. The written informed consents were obtained from each participant prior to enrollment. In this experiment, we obtained acceleration and gyro values at the same time, but confirmed

that the gyro values represent the classification accuracy and characteristics of the gesture more accurately, and only the triaxial gyro value was used in the experiment

The relationship between x, y, and z axis gyro sensor data and gesture is shown as a pairwise correlation graph in order to characterize classification according to acquired gesture. The pairwise correlation graphs were plotted using the seaborn library and are shown in Fig. 5. The following graph shows that the gyro sensor data according to the state is divided into groups according to each gesture.

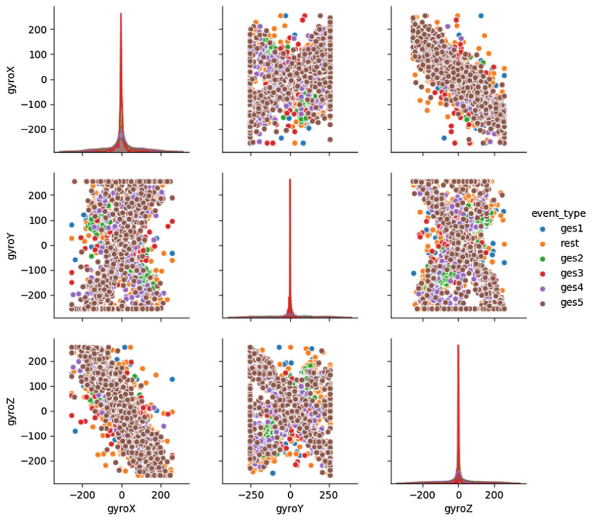


Fig. 5. In order to prevent the overfitting of the neural network, 10-fold cross validation was applied by dividing the input data into 10 sets

In this experiment, 5 kinds of gestures were added, and the state of no gesture was added to label the data in 6 states. The problem of classifying the following states is a multi-classification problem in which there are six classes in total, and one-hot-encoding the six states invoked from the database to create a neural network made of Tensorflow, a python deep learning framework. The total number of sensor data records for learning was limited to 5346, and the total number of epochs was set to 50000. The input layer of the neural network receives 3 - axis gyro data and sends it to 512 nodes and uses Relu as an activation function. The number of hidden layers is 512, and sigmoid and ReLu are repeatedly applied to each layer as an activation function, and there are 4 hidden layers. The output layer has three outputs and softmax as an activation function. The model of the neural network used is shown in Fig. 6.

The accuracy of cross validation is shown in Fig. 7. The accuracy of the overall performance is 97.178%. In Fig. 8, we set the epoch to 10000 and plot the loss and accuracy for each iteration. As the Epoch increases, the accuracy increases. However, when the number of epochs increases, the accuracy does not change. If the gyro sensor and the acceleration sensor data are received at the same time in the future, accuracy can be further improved.

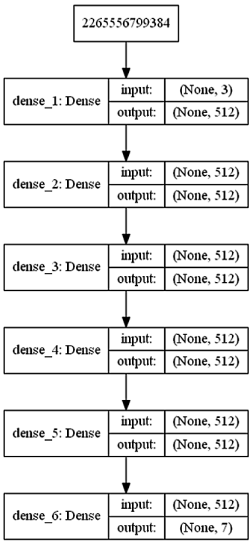


Fig. 6. The configuration of deep neural network models for our experiments

Cross validation	1	2	3	4	5	6	7	8	9	10	Average
Accuracy	97.23	96.07	97.74	96.97	97.59	97.16	97.04	97.61	96.75	97.62	97.178

Fig. 7. Experimental results by cross validations.

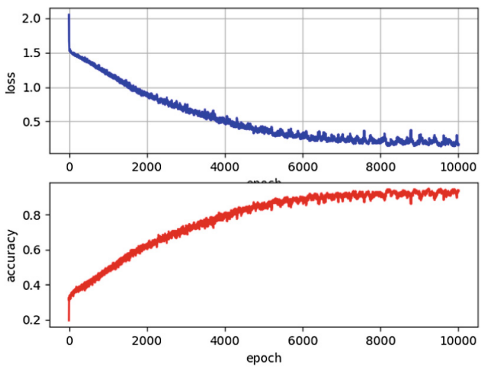


Fig. 8. The gesture recognition accuracy and loss graph by each epochs.

5 Conclusion

The purpose of this study is to improve convenience of production of personal broadcasting by performing functions necessary for broadcasting through gesture recognition using deep learning. For this, we implemented a deep learning based gesture recognition system for immersive broadcasting production that defines the functions to be performed by each gesture. There are two main methods of gesture recognition. A method for image processing and a method using a sensor, Since the method using image processing is affected by the user's camera and the production environment, this paper uses a gesture recognition method using only a sensor. In the future, it will be applied not only to the convenience of production of personal broadcasting through the situation recognition system through deep learning but also to the safety management system of the system and the construction site, which is a fusion of many IoT such as Smart City.

Acknowledgement. This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No. 2016-0-00099, Personal Broadcast Technology Development for Production Convenience and Maximum Viewing Experience).

References

1. Choi, J., Lee, H., Lee, S.: Deep learning-based hand gesture recognition algorithm using multi-modality information. In: 2017 Summer Conference, pp. 672–673. The Institute of Electronics and Information Engineers (2017)
2. Sood, S.K., Mahajan, I.: Wearable IoT sensor based healthcare system for identifying and controlling chikungunya virus. Comput. Ind. **91**, 33–44 (2017)
3. Hiremath, S., Yang, G., & Mankodiya, K.: Wearable Internet of Things: concept, architectural components and promises for person-centered healthcare. In: 2014 EAI 4th International Conference on Wireless Mobile Communication and Healthcare (Mobihealth), pp. 304–307. IEEE, November 2014
4. Varatharajan, R., Manogaran, G., Priyan, M.K., Sundarasekar, R.: Wearable sensor devices for early detection of Alzheimer disease using dynamic time warping algorithm. Cluster Comput. 1–10 (2017)
5. Castillejo, P., Martinez, J.F., Rodriguez-Molina, J., Cuerva, A.: Integration of wearable devices in a wireless sensor network for an E-health application. IEEE Wirel. Commun. **20**(4), 38–49 (2013)
6. NG, K.K.R., Rajeshwari, K.: Interactive clothes based on IOT using NFC and Mobile Application. In: 2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC), pp. 1–4. IEEE, January 2017
7. Perera, C., Jayaraman, P., Zaslavsky, A., Christen, P., Georgakopoulos, D.: Dynamic configuration of sensors using mobile sensor hub in internet of things paradigm. In: 2013 IEEE Eighth International Conference on Intelligent Sensors, Sensor Networks and Information Processing, pp. 473–478. IEEE, April 2013
8. Zhou, W., Piramuthu, S.: Security/privacy of wearable fitness tracking IoT devices. In: 2014 9th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–5. IEEE, June 2014

9. Rautaray, S.S., Agrawal, A.: Vision based hand gesture recognition for human computer interaction: a survey. Artif. Intell. Rev. **43**(1), 1–54 (2015)
10. Garg, P., Aggarwal, N., Sofat, S.: Vision based hand gesture recognition. World Acad. Sci. Eng. Technol. **49**(1), 972–977 (2009)
11. Lee, S.: Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. Int. J. Remote Sens. **26**(7), 1477–1491 (2005)
12. Fleury, A., Vacher, M., Noury, N.: SVM-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results. IEEE Trans. Inf Technol. Biomed. **14**(2), 274–283 (2010)

Design and Implementation of an Efficient Web Crawling Using Neural Network

Ahmed Md. Tanvir, Yonghoon Kim, and Mokdong Chung(✉)

Department of Computer Engineering,
Pukyong National University, Busan, Korea
tva.csai@gmail.com, kimyhjava@pukyong.ac.kr,
mdchung@pknu.ac.kr

Abstract. The number of users the usage of internet is mounting day by day. Currently, researches on the information using the retrieval model neural networks have been actively progressed for the retrieval of information and the classification of documents. Various types of algorithms have been applied for identification and quantification of the words weights in documents. As information technologies accelerate, it is necessary to understand the exact meaning of documents through analyzing the words, using the advanced methods of technologies. In this paper, specific keywords were used by word2vec to identify naturally fused word frequencies, semantic relationships, and directional text-ranks. Therefore, the neural network is the advanced mechanism to verify the semantic relationship between words and texts in a particular document. Our approach uses the Word2vec to capture the semantic features between words in the selected text, and meanwhile naturally integrate the word frequency, semantic relation.

Keywords: Web crawler · Word2vec · Hyperlink · Reinforcement learning · Q-learning

1 Introduction

In recent years, the widespread use of the internet is increased. As a result, the crawler research system is improving every day because there is no alternative to the crawler to make the internet more dynamic and easier for the use of the user. In order to continue this aptitude of improvement, we research the semantic features in this paper. Earlier, in the researches of semantic features, they researched a particular information, where they describe the 'semantic features' in a condensed way and their accuracy rate is not up to the satisfaction levels [3, 5]. In this proposed system, we focused on the semantic features between words in the text using word2vec [2]. Word2vec is considered one of the easiest and popular algorithms to identify the relationship or quantify weight among the documents. In this paper, firstly we have chosen sky, love, life, university and war as specific keywords from the internet and used the Reinforcement Learning (RL) or Q-learning [4] algorithm by improving the equation (i.e. 1/δ) to find out the semantic relation of these words in different documents. The RL or Q-learning is known to learn unrealistically high action values because it includes a maximization step over

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 116–122, 2020.
https://doi.org/10.1007/978-981-13-9341-9_20

estimated auction values, which tends to prefer overestimated to underestimate values. Therefore, in this paper, we have used the RL algorithm. The rest of the paper has been constructed in the following manner. The next section presents the related work. In Sect. 3, our implementation and evaluation of the system. Finally, Sect. 4 discusses the conclusion and future work.

2 Related Work

2.1 Web Crawling Background

The first Internet "search engine", a tool called "Archie" shortened from "Archives", was developed in 1990 and downloaded the directory listings from specified public anonymous FTP (File Transfer Protocol) sites into local files, around once a month [1]. The introduction of the World Wide Web in 1991 has numerous of these Gopher sites changed to websites that were properly linked by Hypertext Markup Language (HTML) links. At that time "Gopher" was created, that indexed plain text documents. In the year 1994, "WebCrawler" was launched the first "full text" crawler [3] and search engine. In early 2000, crawler researchers made some changes in their research. They tried to improve the quality of web search engines. For this reason, they split the web search engine through "Hypertext" and "Hyperlink" [5]. More changes in web crawler were achieved in 2010. The crawler or web search engine's "hypertext" and "hyperlink" are described as "data web mining", "focus web", and so on. The web crawler or web search engine's "hypertext" and "hyperlink" are divided by "data mining", "focus web", where 'deep web crawler' and 'web crawler hidden' data are described.

2.2 System Overview

In Fig. 1, illustrates the system overview. The working of the system overview discussed as follows:

1. Selecting a starting seed URL or URLs.
2. Word separation using word2vec.
3. Choosing the word with the highest weight using word2vec, if not success then go back to the previous step.
4. If the words are a successful selection then it goes to the next step, otherwise it will be back before.
5. Verify the semantic relation between the words and texts in a particular document using the RL algorithm.
6. If the semantic relation is able to detect then the work is success. If not, then introduce escape and avoidance response.

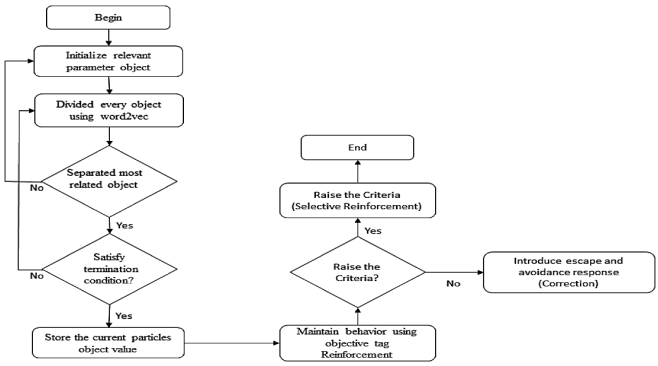


Fig. 1. Working system overview

2.3 Reinforcement Learning

Reinforcement Learning or Q-Learning is learning method, which renews each synapse weights according to output signals. The learning rule can be applied to feed forward networks and recurrent networks. The flow of measurement of the renewal of the syndrome is shown by the pseudocode as follows.

1. Initialize state (x_n, a_n)
2. Calculate the reward of action an
3. for each document $d_i \in R(x_{n-1}$, an)
4. for each keyword K in di do
5. if action a (K) $\neq$ A then A = A U a (K)
6. else then restart Reinforcement-learning (RL) of action a (K)
7. end for
8. end for
9. Change the current state to x_{n+1}
10. $P_r = P_r$ U $[a_n]$; restart candidate set S; S = S / $[a_n]$
11. for each $a_i \in$ S update its reward r
12. for each $a_i \in$ S update its each value δ
13. for each $a_i \in$ S calculated its RL-value
14. return min $\max_a [Q(x_n, a)]$
15. End

Specifically, the surfacing algorithm first calculates the reward of the last executed action and then updates the action set through Step 3 to Step 8, which motive the agent to transit from its state 'x_n' to the successor state 's_{n+1}'. Then the training and candidate set are restart in accord with the new action set in Step 10. After that, the algorithm catalog the reward and RL-value for each action in candidate set in Step 11, Step 12 and Step 13 respectively. The action that maximizes RL-value will be returned as the next to be executed action.

3 Implementation

In order to prove that the accuracy of the semantic analysis can be improved by applying web data, two experiments are conducted in this paper. This of one is an experiment to distinguish document features by extracting features of document semantics and another one is Word Embedding [3], which is used to distinguish the meaning of words in a document. The data used to test that is an online website corpus like NASA, Korean University and so on related to Sky, Love, Life, War and University and four data sets are created to check the performance of web devices by dividing mesne the data. We chose the closest data from the total data of the keyword using word2vec in Fig. 2.

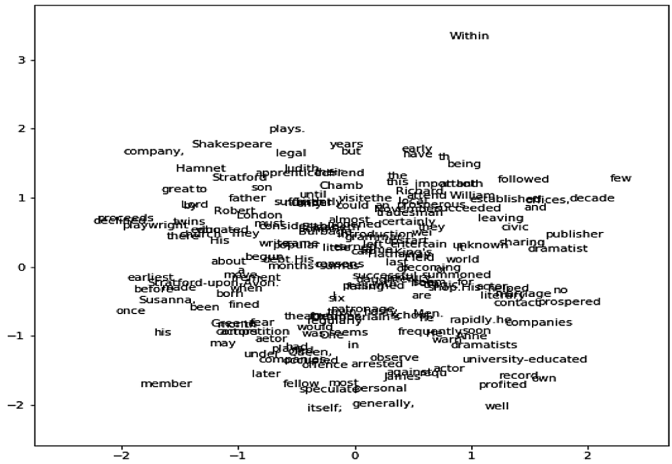


Fig. 2. Word embedding using Word2vec

Reinforcement Learning is a learning method, which renews each synapse weights according to output. The learning rule can be applied to feed forward networks and recurrent networks. The calculated flow of renewing synapse weights is shown in follows. Firstly, the influence of rewards 'r' and the influence of document 'δ' are obtained by

$$r = \sum_{i=0}^{n} \gamma^{i-1} r_{\mathrm{i}} \tag{1}$$

$$\delta = (1/\mathrm{r})^{(1/1+\mathrm{r})} \tag{2}$$

$$\mathrm{Q(s,a)} \leftarrow \mathrm{Q(s,a)} + \mathrm{a}/\delta\,[\mathrm{r} + \gamma \cdot \max_{\mathrm{a}} \mathrm{Q}(\mathrm{s_{n+1}}, \mathrm{a}') - \mathrm{Q}(\mathrm{s}', \mathrm{a}')] \tag{3}$$

$$\mathrm{Q(s,a)} \leftarrow \mathrm{Q(s,a)} + \mathrm{a}\,[\mathrm{r} + \gamma \cdot \max_{\mathrm{a}} \mathrm{Q}(\mathrm{s_{n+1}}, \mathrm{a}') - \mathrm{Q}(\mathrm{s}', \mathrm{a}')] \tag{4}$$

Following the aforementioned equation, we compute the reward and value of each corpus. Here we can see in Fig. 4 that the accuracy rate is 75%, because the baseline system selected the document value randomly using in Eq. (4). In order to solve this problem we multiply the value "1/δ" using in Eq. (3). As a result, our accuracy rate has increased by 20%. Therefore, our accuracy rate is 95% in Fig. 3. Here we have used twenty document, which can be seen as "True" and "False" in Figs. 3 and 4. These results show that if the same word appears in every document, it excludes words with a value of "False", which did not produce good results in learning. So accuracy based on learning outcomes was unsatisfactory. On the other hand, the results show that if the same word does not appear in every document, it excludes words with a value of "True", which provides good results. So accuracy based on learning outcomes was satisfactory.

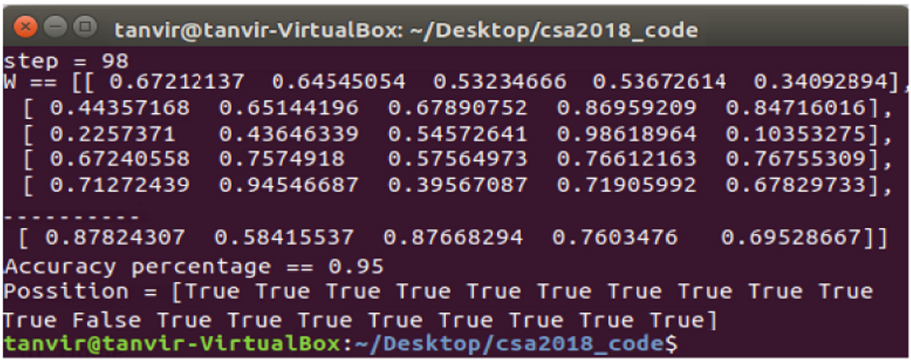


Fig. 3. Proposed system 1/δ using result

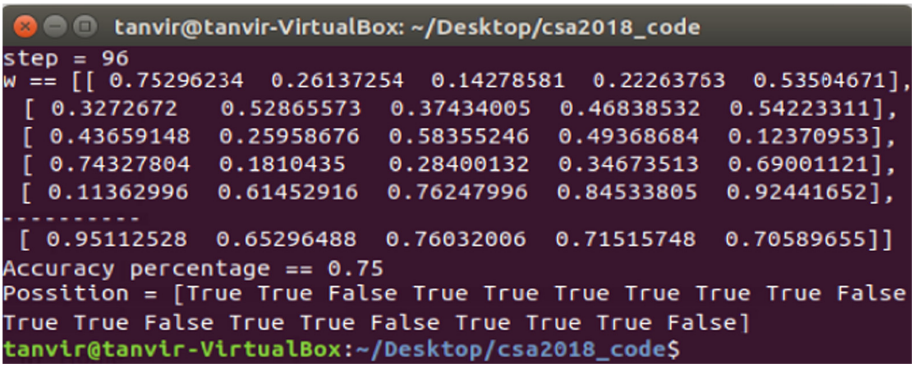


Fig. 4. Existing system result (not use 1/δ)

3.1 Evaluation

How to evaluate the web data in crawling it is a difficult question. Generally, there have two kinds of evaluation methods: intrinsic evaluation and extrinsic evaluation. Intrinsic evaluation directly analyzes the data to judge the quality of themselves. Extrinsic evaluation judges the quality of data by its affection to some other task. In this paper, an intrinsic evaluation method is presented. In this paper, we applied the word2vec to capture the semantic features between words in text. However, we used the RL algorithm to calculate the semantic features of the document. Where '1/δ' is used to find out the value of the words separately. In the baseline method, the word value is randomly captured, in this reason, their total word value is not clear and the accuracy rate is lower. However, using '1/δ' in the proposed system increased accuracy by 20%. In Table 1 we described the properties before and after using '1/δ' at RL-algorithm flowing Eqs. (3) and (4).

Table 1. Each document separate value

	Love	War	University	Sky	Life
Before using δ function	0.513531	0.441742	0.421122	0.491121	0.514111
After using δ function	**0.661431**	**0.723112**	**0.752131**	**0.644142**	**0.756521**

4 Conclusion

This paper, the main topic of semantic relation based on Reinforcement Learning (RL) is studied and the influence of different word embedding on keyword accuracy is compared. In this paper, we use two methods to reach our goal, one is Word2vec and the other is RL algorithm, that is, we use Word2vec to convert text to word vector and improve the RL algorithm from two aspects, and finally, we extract semantic relation with the improved method. Our experimental results are compared with the five data sets. Experimental results suggest that our algorithm can show the semantic features of the web pages very quickly and properly. For future work, we plan to extend our work to consider additional crawling Web-services in our validation to generalize our findings. In addition, we are also planning to extend our approach to balance between refactoring in the code and interface levels.

Acknowledgements. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF2017R1D1A1B03030033).

References

1. Amudha, S., Phil, M.: Web crawler for mining web data. Int. Res. J. Eng. Technol. (IRJET) **4** (2017)
2. Kim, Y., Hong, H., Chung, M.: Application of cohesion devices for improvement of distributional representation. In: Proceeding of the 14th International Conference on Multimedia Information Technology and Applications (MITA 2018), Shanghai University of Engineering Science, China, 28–30 June 2018, pp. 84–87 (2018)
3. Zhao, D., Du, N., Zhi, C., Li, Y.: Keyword extraction for social media short text. In: 14th Web Information System and Applications Conference (WISA), pp. 251–256 (2017)
4. Jiang, L., Wu, Z., Feng, Q., Liu, J., Zheng, Q.: Efficient deep web crawling using reinforcement learning. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 428–439. Springer, Heidelberg (2010)
5. Patil, Y., Patil, S.: Implementation of enhanced web crawler for deep-web interfaces. Int. Res. J. Eng. Technol. (IRJET) **3** (2016)

An Intuitive VR Interaction Method for Immersive Virtual Reality

Ruixin Zhao[1], Namjung Kim[2(✉)], and Kyoungju Park[3]

[1] Department of Image, Graduate School of Advanced Imaging Science, Multimedia and Film, Chung-Ang University, Seoul, Korea
qustzrx@outlook.com
[2] Chung-Ang University Industry Academic Cooperation Foundation, Seoul, Korea
sogalman@naver.com
[3] Department of Software, Chung-Ang University, Seoul, Korea
kjpark@cau.ac.kr

Abstract. We present an intuitive virtual reality interaction method that manipulates virtual objects and expresses hand gestures for immersive virtual environment. The proposed object manipulation model includes virtual reality interactions such as grabbing, holding, dropping, and shaking, where haptic feedbacks corresponding to user's hand movements are applied. We apply acceleration based on a finite difference method to generate realistic vibrations for haptic interaction. Our system also incorporate a efficient method to visualize gestures in an intuitive way for users to change the gesture by pressing the controller's touchpad. The experimental results show a virtual reality interaction in which a user manipulates objects and expresses hand gestures in a virtual environment.

Keywords: VR interaction · Object manipulation · Hand gesture · Haptic feedback · Immersive VR

1 Introduction

Manipulating objects and expressing gestures is daily found in VR games as well as in the real world. Virtual reality Interaction researches [1] are actively conducted in the field of computer graphics and virtual reality to narrow the gap between the real world and the virtual world. Also, the demand for immersive virtual reality applications is rapidly increasing due to the spread of virtual reality devices such as the Oculus Rift and the HTC Vive. Especially, virtual reality interaction technology, which manipulates virtual objects and expresses gestures in an easy and convenient manner in the immersive environment, is an important factor in virtual reality applications. Our goal is to develop an unified system that allows users to manipulate objects and express hand gestures in a familiar way in virtual reality, just as a user manipulates objects and expresses hand gestures in everyday life.

A considerable number of techniques have been developed to manipulate virtual objects. Ray based selection technique [2] is widely used for object manipulation in the

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 123–128, 2020.
https://doi.org/10.1007/978-981-13-9341-9_21

virtual environment that casts a ray from the user's hand or eye, and selecting the objects colliding with the ray, but this technique involve eye-hand visibility mismatch problem. Argelaguet et al. [3] presented a mapping for ray control, which improve the eye-hand visibility mismatch in cluttered virtual environments. Dam et al. [4] proposed a Kinect based interaction technique that provides convenience to users because it handles user's movements without an object attached to the human body, but it is difficult to convey a stronger sense of VR immersion to users because haptic feedback and HMD (Head Mounted Display) are not provided. Caputo et al. [5] presented a usability study for object manipulation tasks on a immersive VR setup that is based on the Oculus DK2 for scene display and a Leap Motion controller for gesture interaction, however, this study did not consider how to provide haptic interaction to the user. In this paper, we present a novel virtual reality interaction methods that incorporate object manipulation including haptic feedbacks and hand gestures animation.

In summary, our contributions include: (1) An intuitive interaction method that allows users to easily grab, hold, drop, shake, and haptic feedback on virtual objects. (2) A user-friendly interface that allows the user to press the controller's touch pad with their fingers to express the desired hand gestures and change the hand gestures.

2 Interaction with Objects

In the virtual reality environment, we manipulate the objects including grab, hold, drop and shake etc. Object manipulation is one of the most important influencing factors of interaction. We design some easy but effective methods to manipulate objects and implement vibration. We also try to make it more realistic to improve the users' immersive feeling.

2.1 Object Manipulation

To grab the objects, we add the box collider to our virtual hands and the objects which are designed to be manipulated. When user presses the controller trigger button, we check everything that overlapped into the hand collider, then we set the transform of the object to the overlapped hand transform for grabbing the objects. When user holds the trigger button, we keep getting the hand transform and assignment it to the object transform for holding the objects so that the grabbed object can move with hand. When user releases the trigger button, we get the hand position and velocity, then we apply the gravity to the object to drop it. In this way, the dropped object will simulate physics with the position and velocity we get in the last frame before release trigger button as shown in Fig. 1.

2.2 Vibration Implementation

Tambourine manipulation is almost same to common objects manipulation except the vibration. When we shake the tambourine, we should have the haptic that the tambourine is shaking. To make the vibration more realistic, we get the velocity of the tambourine in

each frame and calculate the acceleration using finite difference method [6]. acc $= \frac{|v-v_p|}{t_\delta}$, v is the velocity in the current frame. v_p is the velocity in the previous frame. t_δ is the time between two frames. If the acceleration that we get is bigger than the value1, we make the controller shake. At the same time, when we shake the tambourine intensely, user should also feel the vibration stronger from the controller. We limit the strong level not to be infinitely large. If the acceleration is bigger than the value2, we keep the strongest level without raising as shown in Fig. 2.

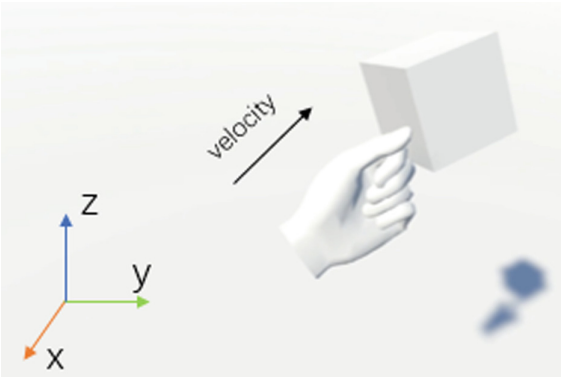


Fig. 1. Get the hand position and velocity in the last frame.

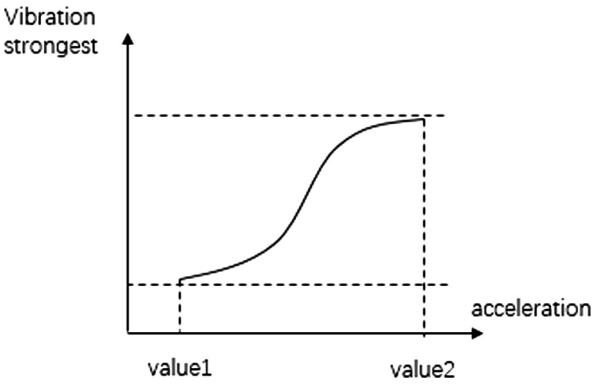


Fig. 2. Relationship between vibration and acceleration.

3 Hand Gestures

To enrich the virtual reality interaction, we design some hand gestures. User can change his hand gesture by pressing the touchpad in the controllers. In this way, user can get more fun from interaction except object manipulation. In our implementation, we designed 6 hand gestures and simply divide the touchpad to 6 parts corresponding

to 6 gestures as shown in Fig. 3. Table 1 shows the method how we divide the touchpad. Then we detect the user's current touch position and display the corresponding gesture animations. When user releases the touchpad button, we clear all the hand gestures to be ready for the other gestures.

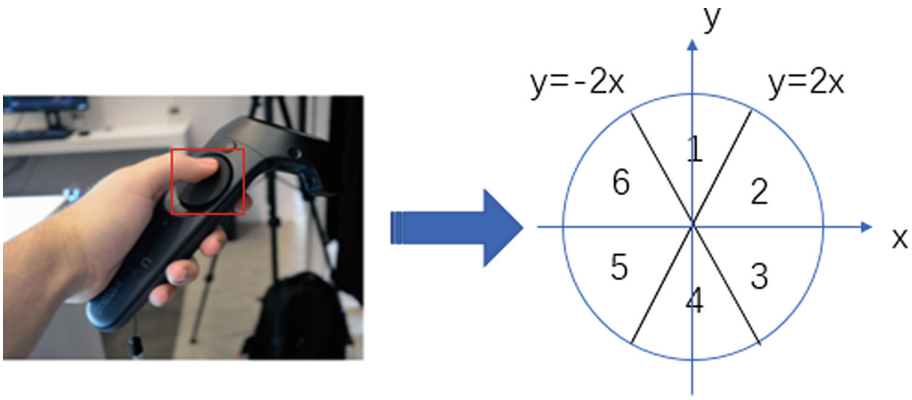


Fig. 3. Correspondence between touchpad region and hand gesture.

Table 1. Touchpad division method.

Hand gesture	Touchpad region
1	$y > 0, y > 2x, y > -2x$
2	$y > 0, y < 2x$
3	$y < 0, y > -2x$
4	$y < 0, y < 2x, y < -2x$
5	$y < 0, y > 2x$
6	$y > 0, y < -2x$

4 Results

The proposed VR interaction system were implemented by using Unity3D 2017.3.0f3 and HTC Vive. Figure 4 shows our objects manipulation result. And by recognizing the different tags of objects, we add the vibration property to tambourine. In the result, we can manipulate objects easily and feel the haptic of vibration naturally. Our hand gestures result is shown in Fig. 5 and different gestures can be switched smoothly. We can express our feelings by showing different hand gestures, which is very interesting and effective in virtual reality interaction.

Fig. 4. Microphone manipulation illustration (top). Tambourine manipulation illustration (bottom).

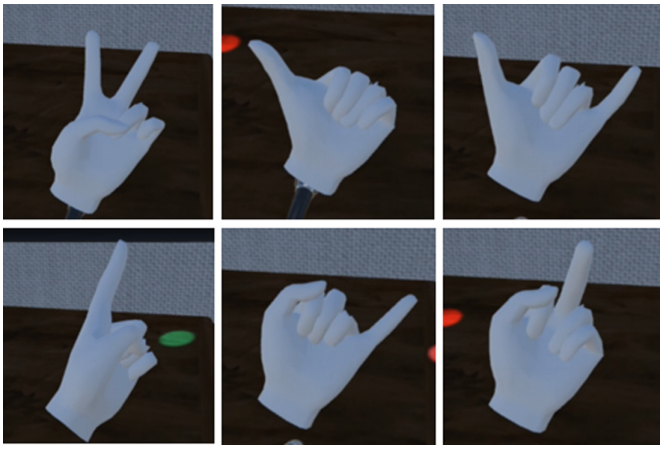


Fig. 5. Hand gestures interactions result.

5 Conclusion

In this paper, we present a novel virtual reality interaction method for immersive virtual environment including object manipulation, tambourine vibration and hand gestures interaction. Our object manipulation technique includes grabbing, holding, dropping,

and shaking. Proposed system provide effective haptic feedbacks corresponding to user's hand shaking by applying acceleration based on a finite difference method. Also our system generate hand gesture animations in an intuitive manner for users to change the gesture by pressing the touchpad. Our methods allow users interact with the virtual reality easily and naturally. Besides, the hand gestures interaction design is very helpful for users to have a better virtual reality experience.

Acknowledgments. This research is supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program 2018, and is supported by the Mid-Career Researcher Program through an NRF grant funded by the MEST (No. NRF-2016R1A2B4016239).

References

1. Marks, R.: VR interactions. In: SIGGRAPH 2017 Courses, Los Angeles, CA, USA, 30 July–03 August 2017 (2017)
2. Steed, A.: Towards a general model for selection in virtual environments. In: IEEE Symposium on 3D User Interfaces 2006 (3DUI 2006), pp. 103–110 (2006)
3. Argelaguet, F., Andujar, C.: Efficient 3D pointing selection in cluttered virtual environments. IEEE Comput. Graph. Appl. **29**(6), 34–43 (2009)
4. Dam, P., Braz, P., Raposo, A.: A study of navigation and selection techniques in virtual environments using microsoft Kinect®. In: Virtual Augmented and Mixed Reality. Designing and Developing Augmented and Virtual Environments, pp. 139–148. Springer, Heidelberg (2013)
5. Caputo, F.M., Giachetti, A.: Evaluation of basic object manipulation modes for low-cost immersive virtual reality. In: 11th Biannual Conference on Italian SIGCHI Chapter, pp. 74–77. ACM (2015)
6. Janabi-Sharifi, F., Hayward, V., Chen, C.-S.J.: Discrete-time adaptive windowing for velocity estimation. IEEE Trans. Control Syst. Technol. **8**, 1003–1009 (2000)

Design and Implementation of a Partial Denoising Boundary Matching System Using Indexing Techniques

Bum-Soo Kim(✉) and Jin-Uk Kim

Department of Future Technology and Convergence Research, Korea Institute of Civil Engineering and Building Technology, Goyang, Republic of Korea
{bumsookim, jukim}@kict.re.kr

Abstract. In this paper we design and implement a partial denoising boundary matching system using indexing techniques. Converting boundary images to time-series makes it feasible to perform fast search using indexes even on a very large image database. Thus, using this converting method we develop a client-server system based on the previous partial denoising research in the GUI environment. The client first converts a query image given by a user to a time-series and sends denoising parameters and the tolerance with this time-series to the server. The server identifies similar images from the index by evaluating a range query, which is constructed using inputs given from the client, and sends the resulting images to the client. Experimental results show that our system provides much intuitive and accurate matching result.

Keywords: Boundary matching · Time-series matching · Partial denoising

1 Introduction

Recent exponential growth in data sizes has resulted in an urgent need for methods that can find useful information quickly in a database. In addition, Efforts [1–4] have recently been made to solve the boundary matching problem for a large database. Converting boundary images to time-series makes it feasible to perform boundary matching even on a very large image database. A *time-series* is a sequence of real numbers representing values at specific time points. Boundary matching first converts boundary images to time-series, called *boundary time-series*, and identifies *data time-series* similar to a given *query time-series*.

In this paper we deal with the design and implementation of an index-based boundary matching system supporting partial denoising. The previous work [4] addresses boundary matching considering partial noise, which is a limited amount of noise embedded in a boundary image. The partial noise varies with level, position, and length on a boundary image. In the case of query images, partial denoising is simple as it is performed only once by prepressing. On the other hand, in the case of data images, it is a challenging problem since all possible partial noises from the data images have to be considered. To solve this problem, they proposed *partial denoising boundary matching*. This matching, however, incurs disk I/O overhead since all the data images are accessed at least once.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 129–134, 2020.
https://doi.org/10.1007/978-981-13-9341-9_22

To solve the disk I/O overhead on partial denoising boundary matching, we first propose an index-based approach that uses a multidimensional index. To build the index, after transforming the partial denoising time-series to low-dimensional points each data time-series, we construct a low-dimensional minimum bounding rectangle (MBR) that contains these points and then store the constructed MBR in a multidimensional index such as an R*-tree [5]. We can construct the MBR since the entries between partial denoising time-series are very similar; that is, even when the partial denoising position is different, many original entries in the partial denoising time-series are frequently used in other partial denoising time-series.

We design and implement a client-server system on GUI environment for users' convenience. The system first constructs the boundary database from data time-series converted from boundary images. It also exploits a low-dimensional transform and a multidimensional index for fast matching. The client converts a query image given by a user to a query time-series, and it then sends a tolerance ε as the similarity measure and a denoising level d, a denoising length l with the query time-series to the server. The server gets similar boundary images by searching the boundary database using the query time-series, the given tolerance ε, and the denoising parameters, and it then sends those similar images to the client. Through experiments, we show that our system provides much accurate and intuitive matching by supporting partial denoising. On the basis of these results, we believe that our index-based system is a practical and interactive way of realizing partial denoising boundary matching.

2 Related Work

2.1 Time-Series Matching

Time-series matching is a problem of finding *similar* time-series to a given time-series in a time-series database [1, 6, 7]. We say two time-series X and Y are similar if the distance $D(X, Y)$ is less than or equal to the user specified tolerance ε [1, 2, 6, 7]. In this paper, we focus on the simplest similarity model that uses the Euclidean distance and the ε-based range query [1–3]. Time-series matching methods can be classified into two categories [7]: *whole matching*, where the lengths of the data time-series and the query time-series are all identical, and *subsequence matching*, where the lengths of the data time-series and the query time-series are different.

Index-based time-series matching also classified into these two categories. Whole matching proposed by Agrawal et al. [6] first stores an MBR containing low-dimensional points into an index, after transforming data time-series into low-dimensional points, and then finds data time-series similar to the query time-series on the index. Subsequence matching proposed by Faloutsos et al. [7] is a generalization of whole matching. Subsequence matching first divide data time-series into sliding windows and store MBRs constructed using their low-dimensional points in an index. Finally, they obtain subsequences similar to the query time-series on the index. In this paper, we focus on the application of whole matching to boundary matching.

2.2 Image Matching

Image matching, also known as content-based image retrieval (CBIR), is a problem of finding data images similar to a given query image by using features of an image [8–10]. The application areas of image matching differs depending on which features are used. In this paper, we focus on shape features, which are useful when color or texture features of images have similar values [9], for image matching. We use the centroid contour distance (CCD) function [1, 2, 11], which is the simplest method for representing one-dimensional shapes, as a shape feature to exploit time-series matching techniques instead of shape features such as edge, curvature, and region.

The denoising problem discussed in this paper is introduced by Kim et al. [2, 4]. They first deal with *whole denoising boundary matching* that is boundary matching that supports noise removal throughout the boundary time-series for a large image database [2]. They also deal with the denoising problem, i.e., partial denoising boundary matching, in boundary matching by extending *whole denoising* to support *partial denoising* [4]. Whole denoising boundary matching [2] utilizes an index-based matching method for a large image database. On the other hand, partial denoising boundary matching [4] utilizes an efficient matching method that exploits the lower bounding techniques in time-series matching but do not consider indexing for a large image database; that is, the focus is on matching techniques rather than the disk I/O overhead problem.

3 Design of an Index-Based System for Partial Denoising Boundary Matching

In general, traditional indexing methods first transform high-dimensional points into low-dimensional points and then store each of the low-dimensional pints in an index [6, 12, 13]. These indexing methods, however, is not practical since it incurs substantial overhead to store the low-dimensional points transformed from all partial denoising time-series. In a large image database, this problem is very serious. Consequently, an efficient indexing method is needed for index-based partial denoising boundary matching.

The naïve index-based matching system consists of preprocessing, range search, and post-processing steps. In the preprocessing step, we remove partial noises on data time-series by using the moving average transform [2] and then generate *partial denoising time-series*, where is replaced by the noise-removed subsequence instead of its subsequence. As mentioned in previous work [14], each partial denoising time-series of length n is transformed into an f-dimensional point ($f << n$, we call it *low-dimensional transformation*), and the transformed points are stored in an f-dimensional R*-tree [2, 5]. In the range search step, a query time-series is similarly transformed into an f-dimensional point, and a range query is constructed using that point and the tolerance ε. By evaluating the range query on the index, the candidates that are potentially similar to the query time-series are identified. The low-dimensional transformation guarantees there is no *false dismissal*, but may cause *false alarms* [2, 6, 7]. Thus, in the post-processing step, for each candidate time-series obtained, the actual

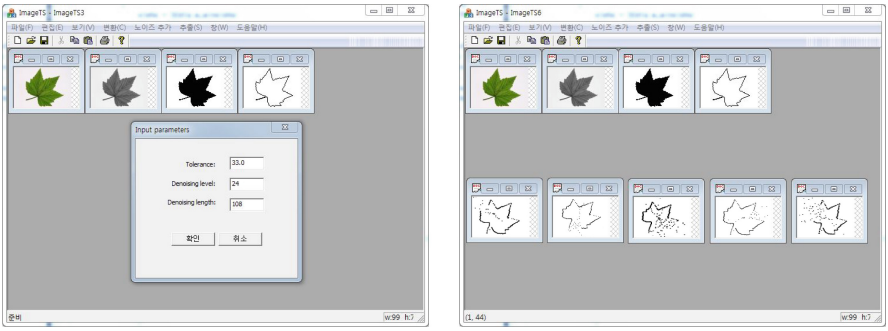
data time-series is accessed from the disk; the distance from the query time-series is computed; and the candidate is discarded if it is a false alarm [2].

To solve the overhead problem, we do not store the *f*-dimensional points but *f*-dimensional MBR [2]. All entries except for the noise-removed subsequence in a partial denoising time-series are frequently used in other partial denoising time-series, even when the denoising position is different; that is, partial denoising time-series generated from one original boundary time-series are similar to each other. This observation enables the indexing process to use an MBR that bounds all partial denoising time-series instead of each partial denoising time-series in an index.

4 Implementation of an Index-Based System for Partial Denoising Boundary Matching

In our system, converting an image to a boundary time-series can be done by two steps: (1) the binary transformation of a gray-scaled image and (2) the boundary tracing of a binary image. Using the binary transformation, we first convert a color image to a gray-scaled image and then obtain a binary image from the gray-scaled image. In this process, we need to determine the binary threshold, which controls the degree of binary transformation [2]. In the experiments, we performed evaluations and chose 240 as the binary threshold. For the boundary tracing of a binary image, we used the well-known tracing algorithm exploiting 8-neighborhood connectivity [10].

We implemented a client-server system for index-based partial denoising boundary matching. Figure 1 shows the input and output screenshots of our system. In Fig. 1(a), we send a query image to the server together with a tolerance and denoising parameters to get the result of partial denoising boundary matching. Figure 1(b) shows the corresponding output images that are similar to the given query image.



(a) An input screenshot of the matching system. (b) An output screenshot of the matching system.

Fig. 1. Example screenshots of the index-based partial denoising boundary matching system.

In the experiments, we used the synthetic boundary dataset used from the previous work [4]. This dataset consisted of 102,590 boundary time-series of length 360 including nine different partial noises created by changing the length and the position

from each original image, all collected from the Web. To generate the partial noise, we used the Gaussian noise model [10]. Although we only used ten thousand original images [1–4], we actually extracted more boundaries than that number using the CCD method since one image might contain multiple boundary objects.

We performed the experiments on the following server. The hardware platform was an IBM compatible PC equipped with a 2.0 GHz Intel Core 2 Duo CPU, 2.0 GB RAM, and 500 GB hard disk. The software platform was the CentOS 6.3. We used the C/C++ language to implement the index-building and matching algorithms. As the multidimensional index, we used the R*-tree and set its index and data page sizes to 4,096 bytes. In addition, we used PAA [1] as the low-dimensional transformation and extracted 72 features each time-series using it.

5 Conclusions

In this paper, we design and implement an index-based system for partial denoising boundary matching. To solve a performance overhead problem for storing boundary time-series in a multidimensional index, we exploit time-series matching techniques. We implement the system based on the client-server model that makes it feasible to support multi-users on a large database in real time. In this paper, we focus on the development of the GUI functions that it is easy to identify and intuitively understand the matching results. As the future work, we will try to integrate our design and implementation supporting *k*-NN queries instead of range queries.

References

1. Moon, Y.-S., Kim, B.-S., Kim, M.S., Whang, K.-Y.: Scaling-invariant boundary image matching using time-series matching techniques. Data Knowl. Eng. **69**(10), 1022–1042 (2010)
2. Kim, B.-S., Moon, Y.-S., Choi, M.-J., Kim, J.: Interactive noise-controlled boundary image matching using the time-series moving average transform. Multimed. Tools Appl. **72**, 2543–2571 (2014)
3. Loh, W.-K., Kim, S.-P., Hong, S.-K., Moon, Y.-S.: Envelope-based boundary image matching for smart devices under arbitrary rotations. Multimedia Syst. **21**(1), 29–47 (2015)
4. Kim, B.-S., Moon, Y.-S., Lee, J.-G.: Boundary image matching supporting partial denoising using time-series matching techniques. Multimed. Tools Appl. **76**, 8471–8496 (2017)
5. Beckmann, N., Kriegel, H.-P., Schneider, R., Seeger, B.: The R*-tree: an efficient and robust access method for points and rectangles. In: The ACM SIGMOD International Conference on Management of Data, Atlantic City, New Jersey, pp. 322–331 (1990)
6. Agrawal, R., Faloutsos, C., Swami, A.: Efficient similarity search in sequence databases. In: The 4th International Conference on Foundations of Data Organization and Algorithms, Chicago, Illinois, pp. 69–84 (1993)
7. Faloutsos, C., Ranganathan, M., Manolopoulos, Y.: Fast subsequence matching in time-series databases. In: The ACM SIGMOD International Conference on Management of Data, Minneapolis, Minnesota, pp. 419–429 (1994)

8. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. **24**(4), 509–522 (2002)
9. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Image retrieval: ideas, influences, and trends of the new age. ACM Comput. Surv. **40**(2), 34–94 (2008)
10. Sonka, M., Hlavac, V., Boyle, R.: Image Processing, Analysis, and Machine Vision. Cengage Learning (2014)
11. Zang, D., Lu, G.: Review of shape representation and description techniques. Pattern Recogn. **37**(1), 1–19 (2004)
12. Berchtold, S., Bohm, C., Kriegel, H.-P.: The pyramid-technique: towards breaking the curse of dimensionality. In: The ACM SIGMOD International Conference on Management of Data, Seattle, Washington, pp. 142–153 (1998)
13. Arandjelovic, R., Zisserman, A.: Three things everyone should know to improve object retrieval. In: The IEEE Conference on Computer Vision and Pattern Recognition, Providence, Rhode Island, pp. 2911–2918 (2012)
14. Kim, B.-S., Moon, Y.-S., Kim, J.: Noise-control boundary image matching using time-series moving average transform. In: The 19th International Conference on Database and Expert Systems Applications, Turin, Italy, pp. 362–375 (2008)

Singing Lip Sync Animation System Using Audio Spectrum

Namjung Kim[1(✉)] and Kyoungju Park[2]

[1] Chung-Ang University Industry Academic Cooperation Foundation, Seoul, Korea
sogalman@naver.com
[2] Department of Software, Chung-Ang University, Seoul, Korea
kjpark@cau.ac.kr

Abstract. When humans sing, the changes in pitch and volume are usually significantly greater than when they speak. The shape and size of the mouth changes dramatically when singing compared with when talking, depending on the strength of the voice. In this study, we propose a new model that effectively generates the change in mouth shape in the vocal environment based on the voice strength estimated by analyzing the audio spectrum. We estimate the voice strength by analyzing the spectrum of the voice using a Fast Fourier Transform-based numerical technique for each frame. We apply the intensity of the estimated voice to the morph targets associated with the mouth shape as the weight of the blendshape. Experimental results show a visually convincing lip-synching animation that changes the mouth shape significantly depending on the pitch and volume of the voice.

Keywords: Singing lip sync animation · Audio spectrum · Blendshape activation · Fast Fourier Transform

1 Introduction

Characters speaking and singing are common in movies, games, and animations. However, when a character's facial animation differs from how a person speaks in the real world, it gives the viewer a sense of discomfort and inconvenience. Therefore, extensive research has been conducted in the computer graphics community to develop a technique for generating lip-sync animation that is visually realistic. With the rapid development of graphics hardware and virtual reality devices over the past decade, there is a growing demand for speeching and singing lip-sync animation technology that can be easily applied to user applications. Although lip-sync animation techniques have shown remarkable progress, most of the existing lip-sync studies have focused on methods by which virtual characters appear to speak realistically. However lip-sync animation for singing character remains a challenging issue.

A considerable number of lip-sync studies have been conducted on techniques for expressing realistic speech animation. Chao et al. [1] presented a fast lip-sync technique for visual speech using a greedy graph search algorithm and Xie et al. [2] presented a mouth animation technique for talking faces that improves visual

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 135–140, 2020.
https://doi.org/10.1007/978-981-13-9341-9_23

articulatory movements using a dynamic Bayesian network. Edwards et al. [3] proposed an expressive lip-sync method that generates a range of speaking styles by using jaw articulation. Recently, deep learning based lip-sync animation techniques have been widely investigated for realistic speech animation. Taylor et al. [4] proposed an effective deep learning approach for generalized speech animation that does not depend on the specific speaker, speaking style, or even the language being spoken. Suwajanakorn et al. [5] presented an efficient video-based lip-sync technique that does not require phoneme labels by using a recurrent neural network, and Karras et al. [6] proposed a realistic 3D facial animation technique that extracts emotional state based on a convolutional neural network architecture. In this study, we propose a singing lip-sync animation method that can be applied effectively in VR application by applying an adaptive mouth shape based on the pitch and volume of voice.

Our major contributions to lip sync animation can be summarized as follows:

- A novel framework that analyzes audio spectrum and automatically generates the singing lip synch animation according to the voice strength.
- A new model that effectively generates the mouth shapes that change variably according to changes in pitch and volume.
- A spectrum based approach that naturally produces the vibrations of the mouth when the character is singing.

2 Spectrum Based Lip Sync Model

In the vocal environment, the shape and size of the mouth varies dramatically depending on the voice strength. We use a spectrum based approach to generate realistic singing lip sync animation. We estimate the strength of the voice by analyzing the voice spectrum in each frame. Then, the estimated voice strength is reflected in the blendshape to smoothly connect the rapid variation in mouth shape. The proposed model can naturally express the variations of the mouth shape.

2.1 Audio Spectrum Analysis

We use FFT (Fast Fourier Transform) to obtain the current audio spectrum at each frame. We adopt the BlackmanHarris FFT window function [7] to accurately compute the current spectrum data, which is given by the following Eq. (1).

$$W[n] = 0.35875 - (0.48829 * cos(1.0 * n/N)) + (0.14128 * cos(2.0 * n/N)) - (0.01168 * cos(3.0 * n/N)) \tag{1}$$

We get the spectrum of the audio computed by the FFT in the array $Spec[n]$, and the length of the array $Spec[n]$ is given by a power of 2, minimum length is 64, maximum length is 8192. As n increases, the precision and smoothness of the spectrum curve increases. Figure 1 shows the variation of the audio spectrum curve over time.

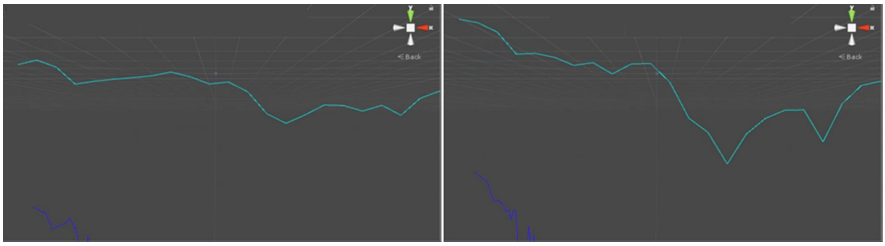


Fig. 1. The variation of the audio spectrum curve over time.

2.2 Audio Strength Estimation

We estimate the strength of voice at each frame with the combination of Gaussian function and audio spectrum array obtained by FFT calculation as shown in Eq. (2).

$$\sum_{i=1}^{n} Spec[i]G(i) \tag{2}$$

Where $G(i)$ denotes a Gaussian function of variance σ as shown in Eq. (3).

$$G(x_i) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x_i^2}{2\sigma^2}} \tag{3}$$

x_i is given by $\frac{i-1}{n-1}$, the variance σ is obtained by squaring the current value x_i minus the mean, dividing it by the total number of elements after adding all the squares. We normalize the acquired voice strength to a range between the minimum and maximum voice strengths of the entire song. The minimum and maximum voice strengths of the entire song are processed in the preprocessing stage.

2.3 Blendshape Activation

The proposed system uses the blendshape method to naturally connect the facial frame of the vocal environment where the mouth shape change is prominent. We apply blendshape weights based on the strength information of the voices to the morph targets placed on the mouth and teeth. Figure 2 shows the schematic diagram of blendshape activation that applies voice strength to morph targets related to mouth shape. The weight of a blendshape is the percentage of the blendshape that is triggered, and it has a value between 0% and 100%. The blendshape has no effect if the weight is 0%, and the blendshape is fully activated if the weight is 100%. In this way, when the pitch of voice is lower, and the volume is smaller, the mouth shape is expressed relatively small, and when the pitch of voice is higher and the volume is larger, the mouth shape is expressed relatively large.

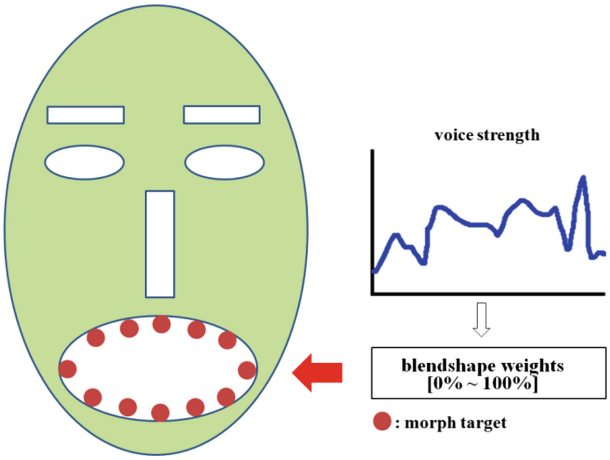


Fig. 2. Schematic diagram of blendshape activation.

3 Results

We implemented our system using Unity3D 2017.3.0f3, C# script, which runs on an Intel i7 4.0 GHZ, 16 GB RAM, Nvidia GTX980 personal computer with FaceFX collapsed XML files. Even though proposed system is currently implemented by using FaceFX collapsed XML data, we believe that our spectrum based approach can be used for other lip sync system. Figure 3 shows the result of a lip-sync animation of the popular ballad song, Queen's love of my life. Figure 4 shows the result of a lip-sync animation of the popular dance song, PSY's kangnam style. Experimental results show that mouth shape changes and vibrations are expressed naturally according to the strength of the voice of the singing character. In addition, both ballad song which have relatively slow and large pitch differences, and dance song which have relatively fast and small pitch differences showed visually convincing results.

Fig. 3. The result of Queen's love of my life.

Fig. 4. The result of Psy's kangnam style.

4 Conclusion

We present a new approach to efficiently generate a singing lip-sync animation by effectively expressing a variable mouth shape based on the audio spectrum. Based on this model, the strength of the voice is effectively estimated by a FFT-based method. The estimated strength of the voice is applied to the morph targets related to the mouth shape as a weight of the blendshape. This creates a mouth shape that smoothly connects between the previous frame and the current frame. The proposed system can naturally animate the large change of the mouth shape of the singing character and the vibration of the mouth shape. The proposed spectrum based method can be implemented easily by extending the existing lip-sync systems.

Acknowledgments. This research is supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program 2018, and is supported by the Mid-Career Researcher Program through an NRF grant funded by the MEST (No. NRF-2016R1A2B4016239).

References

1. Chao, Y., Faloutsos, P., Kohler, E., PIGHIN, F.: Real-time speech motion synthesis from recorded motions. In: Proceedings of ACM SIGGRAPH/Eurographics Symposium on Computer Animation, pp. 347–355 (2004)
2. Xie, L., Liu, Z.-Q.: Realistic mouth-synching for speech-driven talking face using articulatory modeling. IEEE Trans. Multimedia **9**(3), 500–510 (2007)
3. Pif, E., Chris, L., Eugene, F., Karan, S.: JALI: an animatorcentric viseme model for expressive lip synchronization. ACM Trans. Graph. (Proc. SIGGRAPH 2016) **35**(4), 127:1–127:11 (2016)
4. Taylor, S., Kim, T., Yue, Y., Mahler, M., Krahe, J., Rodriguez, A.G., Hodgins, J., Matthews, I.: A deep learning approach for generalized speech animation. ACM Trans. Graph. (Proc. SIGGRAPH 2017) **36**(4), 93 (2017)
5. Suwajanakorn, S., Seitz, S.M., Kemelmacher-Schlizerman, I.: Synthesizing Obama: learning lip sync from audio. ACM Trans. Graph. (Proc. SIGGRAPH 2017) **36**(4), 95 (2017)
6. Karras, T., Aila, T., Laine, S., Herva, A., Lehtinen, J.: Audio-driven facial animation by joint end-to-end learning of pose and emotion. ACM Trans. Graph. (Proc. SIGGRAPH 2017) **36**(4), 94:1–94:12 (2017)
7. Zhang, F.S., Geng, Z.X., Yuan, W.: The algorithm of interpolating windowed FFT for harmonic analysis of power system. IEEE Trans. Power Deliv. **16**(2), 160–164 (2001)

A Tamper-Proof Digital Archiving Scheme Based on Blockchain

Euihyun Jung(✉)

Department of Convergence Software,
Anyang University, Anyang City, Korea
jung@anyang.ac.kr

Abstract. Every society has given a lot of efforts to preserve knowledge for its heritage and the activity is still going on in modern society. The preservation of knowledge in modern society mainly depends on digital technology, unlike previous eras. Digital data is very useful in many respects such as easy transfer or outstanding reusability, but it is not adequate for data preservation because it can be easily modified, vanished, or replaced without notice. In order to resolve this issue, we propose a tamper-proof digital archiving scheme with Blockchain which protects archived data from malicious counterfeiting and guarantees the fidelity of archived data. In the evaluation, the system shows that it is robust to possible forgeries.

Keywords: Blockchain · Immutability · Digital archive · Smart contract

1 Introduction

From ancient times, there have been unceasing efforts to preserve knowledge for delivering human wisdom to descendants. People recorded knowledge onto the material they could easily get from their environments such as stone, papyrus, bamboo, or paper. This kind of effort continues in the Digital Era [1], but the storage medium and access method are completely different from the past times. In the era, the recording of information is not a physical act anymore because most information is digitalized and it is stored in the digital medium. This digitalized recording has advantages over the conventional methods such as easy transfer, easy duplication, less storage space, and outstanding reusability. On the other hand, digitalized information has fatal risks of data consistency. It can disappear without any trace and it can be easily forged without notice.

In recent times, Vinton Cerf pointed out the importance of digital preservation and the necessity of proper methods for the preservation [2]. He sighs that the older techniques such as printing or even microfilm are still used to archive digital information. It seems ironical to use analog media for the digital preservation, but the analog media has a strong point in guaranteeing fidelity.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2018R1D1A1B07049930).

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 141–146, 2020.
https://doi.org/10.1007/978-981-13-9341-9_24

Currently, the main concern of digital preservation is the interoperability of digital archives. If it is developed, once preserved digital archive may be used to make or compose other digital archives. For this purpose, some standards have been announced such as Open Archival Information System (OAIS) [3] reference model and Digital Object Architecture [4] by the Corporation for National Research Initiatives. Comparing to this, anti-counterfeiting for digital archive has been hardly considered yet. However, since digital information can be easily replaced, vanished or forged, the protection method against malicious modifications of digital archives should be studied urgently.

Most people consider WWW is already a huge digital archive for human knowledge, but the WWW content is easily vanished and changed without notice. That means, the WWW content itself is not long-lasting archive ever, but rather the candidate for archiving. For the digital preservation for WWW, there are already Internet Archive [5] service and Archive.is [6] service. The Internet Archive is a San Francisco-based nonprofit digital library with the stated mission of "universal access to all knowledge." This purpose of the Internet Archive is archiving human digital heritages including the WWW content. Archive.is provides people with archiving their requested web pages and retrieving the archived pages. Both services have been doing their good job, but they do not mention about the fidelity of their archiving contents.

Actually, mere digital archiving is not enough to guarantee that an archived one is the genuine and unmodified copy of the original contents. People have no other way but just believe their archived contents will not be forged. The trust for anti-counterfeiting is solely based on not proper technology but the goodwill of service providers. Of course, we think the existing service providers do not intentionally modify the archived contents, but trials of hacking are always possible to forge some significant contents because the archived digital content can be used as history or legal proofs.

In this research, to resolve this issue, we propose a tamper-proof digital archiving scheme based on Blockchain which technically prevents tampering with digital archives. It serves users' requests to archive digital contents and it makes a digital fingerprint and stores the fingerprint into Blockchain to guarantee the fidelity of digital archives.

The rest of this paper is organized as follows. Section 2 describes the design consideration and the details of the proposed architecture. In Sect. 3, several evaluations are conducted to show the robustness of the system against tampering and Sect. 4 concludes the paper.

2 Design of Architecture

2.1 Design Considerations

Same User Experience. There are several considerations in designing the proposed system. First of all, the system should provide the same user experience like other web services. Despite using Blockchain, the system should not force users to know or learn

Blockchain. Also, it enables users to request archiving with only URI for the target web page and to retrieve the archived ones easily.

Limit of Blockchain. From the perspective of Blockchain, the system has to provide a bulk data storing method separately. Usually, Blockchain is not adequate to store unstructured huge data such as image. Therefore, it is carefully designed to maintain the consistency of the bulk data in the external repository and the fingerprint of the corresponding data in Blockchain.

2.2 A Structure of the Proposed System

To adopt the design considerations, we have designed the system as shown in Fig. 1. The system consists of *Portal*, *Virtual Browser*, *Capture Manager*, *Archive Mapper*, *FossilChian,* and two repositories. The *Portal* serves frontend web pages which help users to request archiving and retrieving digital contents. The *Virtual Browser* patches the content of a web page on behalf of its users. The *Capture Manager* captures the page content into a captured image which is visually same as an ordinary human user sees the page. It also stores and retrieves the captured images. The *Archive Mapper* manages an information tuple containing the requested URI, the file path of the captured image for the URI, the fingerprint for the image and etc. The *FossilChain* is a kind of Ethereum blockchain which takes charge in storing and retrieving the information tuple from the *Archive Mapper*.

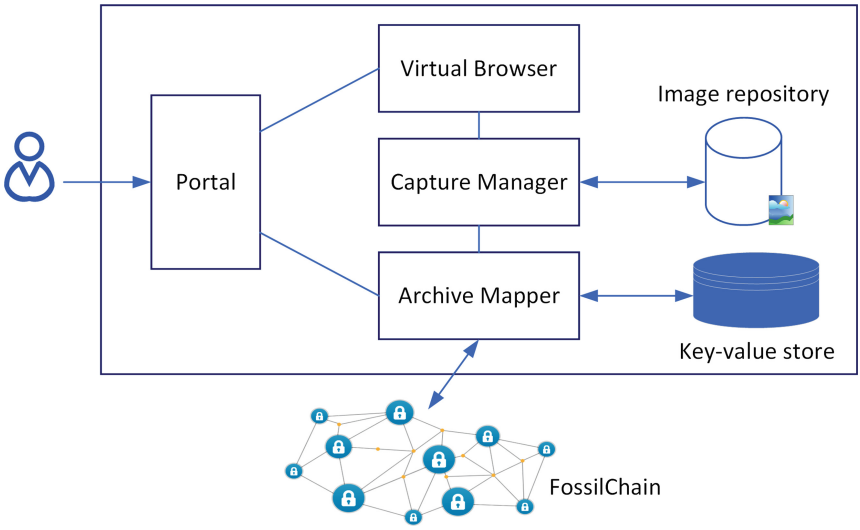


Fig. 1. The structure of the proposed system

2.3 Archiving Operation

In archiving digital contents, the several components of the system cooperate with each other as shown in Fig. 2. When a user requests to archive a web page via the *Portal*, the *Virtual Browser* patches the content of the web page and delivers it to the *Capture Manager*. The *Capture Manager* captures the page content into a capture image and stores the image into the image repository. Then, it generates the hash of the captured imaged and hands over the hash to the *Archive Mapper* with the file path of the stored image file in the image repository. The *Archive Mapper* generates an information tuple containing the archived page's URI, the file location on the repository, the image hash, user information, and timestamp. Then, the *Archive Mapper* requests the *FossilChain* to store the information tuple into an Ethereum blockchain using a smart contract which fossilizes the tuple. After finishing all processes, the *FossilChain* returns a transaction address from the result of the smart contract to the *Archive Mapper*. Then, the *Archive Mapper* generates a shorten URI and stores it with the returned transaction value into the key-value store. After then, the *Archive Mapper* delivers the shorten URI to the *Portal* as a key to access the archived contents.

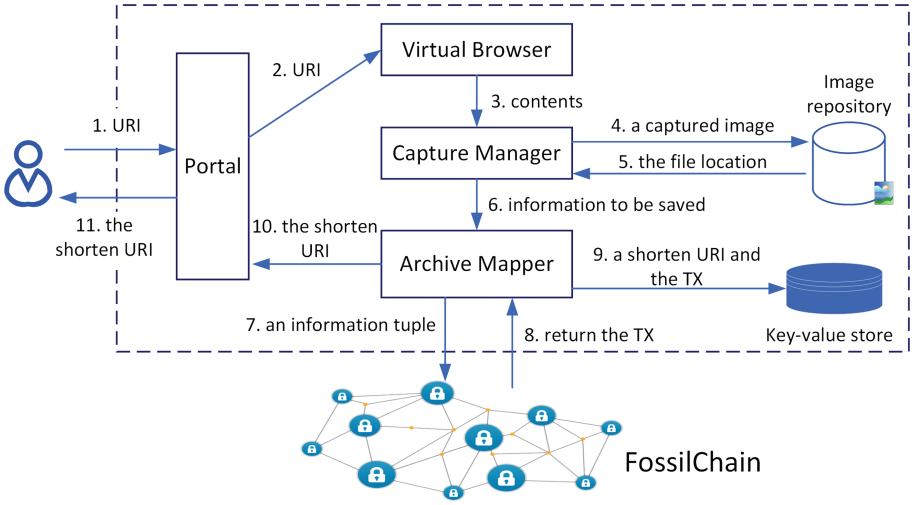


Fig. 2. The operation of archiving digital contents

2.4 Retrieving Operation

If a user wants to retrieve an archived page, the user has to present the page's shortened URI which was given in archiving. On receiving a retrieving request, the *Portal* delivers the request to the *Archive Mapper* to find the linked transaction address from the key-value store. Then, the *Archive Mapper* requests the *FossilChain* to find the information tuple with the found transaction address. After getting the information tuple from the *FossilChain*, the *Archive Mapper* accesses the archived image from the image repository with the file location and calculates a hash from the image. Then, the

Archive Mapper compares the calculated hash with the stored hash in the information tuple to check the fidelity. If the verification succeeds, the *Archive Mapper* delivers the information tuple to the *Portal*, which makes an information retrieval page that enables users to check the archived page.

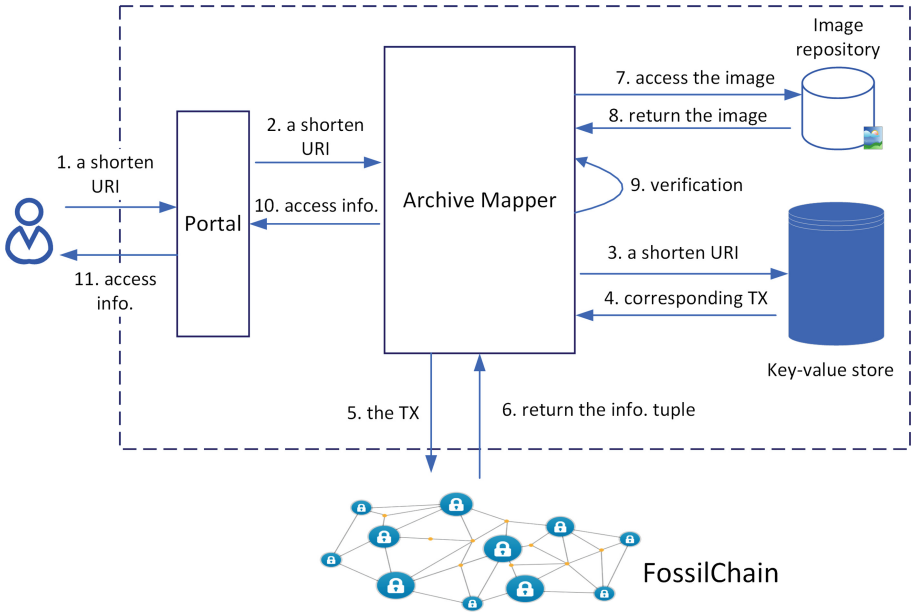


Fig. 3. The operation of retrieving archived digital contents

3 Evaluation

The main check point of the system is whether it can provide tamper-proofing if it is under the risk of hacking. To evaluate this system, we make two security assumptions; First, with proper settings and operations, Blockchain is considered not to be hacked as most experts agree. Second, running processes will not be hacked. Generally, programs before running can be hacked or replaced on a file system, but the running processes are robust to hacking. Therefore, the weak points of the proposed system will be the image repository and the key-value store, because they just maintain static data and they are not designed to be protected from hacking at all.

3.1 Tampering with the Content of the Image Repository

Changing archived images unaware of both users and service providers will be the most critical and possible tampering in digital preservation systems. In the proposed system, if a hacker succeeds to forge the images on the image repository, the forgery will be detected in the retrieving process. During the retrieving process in Fig. 3, the

Archive Mapper calculates a hash dynamically from the archived image from the image repository and compares it with the stored hash from the *FossilChain*. With the verification, the *Archive Mapper* can decide whether the image is forged or not. If a hacker wants to forge an image, the hacker must forge both the target image and the content of Blockchain simultaneously, but we assumed Blockchain will not be hacked.

3.2 Tampering with the Content of the Key-Value Store

The second weak point is the key-value datastore where the system stores a shorten URI and a linked transaction address as a key-value pair. In the system, the shorten URI is used as a key, so the forgery of the shorten URI is meaningless. If a hacker forges a transaction address, it is also meaningless because the transaction address was automatically generated from the execution of a smart contract on the *FossilChain*. The only way to change the transaction address is making a new transaction and replaces the original one with the new one. However, only the *Archive Mapper* can make a transaction and it can be hacked because it is a running process.

4 Conclusion

Every society has given a lot of efforts to record its knowledge as heritages. Although the recording method and medium are changed, the efforts are still going on in the Digital Era. Unlike previous eras, the digital data is the mainstream for conveying and recording knowledge. Digital data has strong points comparing to analog form in many respects; easy transfer, storage capacity, easy duplication, and so on. However, since digital data is easily modified and forged, it is not adequate for persistent data preservation and a new method is necessary to guarantee the fidelity of digital preservation.

In order to resolve this issue, we propose a tamper-proof digital archiving scheme with Blockchain. The system not only provides the same user experience but also protects users' archives from malicious modification. In the evaluation, the system shows it is robust to various kinds of forgeries. Currently, the system has been implemented as a Proof-of-Concept, and we have a plan to evolve the system into a more generalized digital preservation system.

References

1. Dunleavy, P., Margetts, H., Tinkler, J., Bastow, S.: Digital Era Governance: IT Corporations, the State, and E-government. Oxford University Press, Oxford (2006)
2. Cerf, V.G.: The role of archives in digital preservation. Commun. ACM **61**(1), 7 (2017)
3. Lavoie, B.F.: The open archival information system reference model: Introductory guide. Microform Imaging Rev. **33**(2), 68–81 (2004)
4. Sharp, C.: Overview of the Digital Object Architecture (DOA). An Internet Society Information Paper, Internet Society (2016)
5. Internet Archives: About the Internet Archive. https://archive.org/about/. Accessed 21 Sept 2018
6. Archive.is: FAQ. http://archive.is/faq. Accessed 21 Sept 2018

Differential Data Processing for Energy Efficiency of Wireless Sensor Networks

Kwang Kyu Lim[1], JiSu Park[2], Byeong Rae Lee[1], and Jin Gon Shon[1(✉)]

[1] Department of Computer Science, Graduate School, Korea National Open University, Seoul, South Korea
{lahuman, brlee, jgshon}@knou.ac.kr
[2] Division of General Studies, Kyonggi University, Suwon, South Korea
bluejisu@kyonggi.ac.kr

Abstract. Wireless sensor networks use many types of wireless sensors to configure network. However batteries in wireless sensor nodes are energy limited and consume considerable energy for data transmission. Therefore, data merging is used as a means to increase energy efficiency in data transmission. In this paper, we propose Differential Data Processing (DDP), which reduces the size of data transmitted from the sensor node to increase the energy efficiency of the wireless sensor network. Experimental results show that processing the differential temperature data reduces the average data size of the sensor node by 30%.

Keywords: WSN · Data aggregation · Data differential processing · Efficiency · Sensor node

1 Introduction

Wireless sensor networks consist of many types of wireless sensors. Sensor nodes are used for event detection, event ID location detection, and motion control. Sensor nodes are also used to collect data and act as a gateway or router.

However, the location in which the sensor node is used may be an environment where charging or maintaining power is difficult. Given that the lifetime of the sensor nodes has a great effect on the service life of the wireless sensor network, research has been conducted to maximize the senor nodes lifetime by efficiently using their energy [1–8].

The differential data processing for the energy efficiency of the wireless sensor network proposed in this paper provides energy efficient communication by reducing the data size. The reduction is achieved by transmitting the difference values between the previously collected data and the currently collected data.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 147–152, 2020.
https://doi.org/10.1007/978-981-13-9341-9_25

2 LEACH

LEACH (low-energy adaptive clustering hierarchy) is the most prominent protocol in hierarchical routing protocols one that efficiently use cluster-based energy. Here all nodes are clustered and collected, and each cluster has a cluster header. The configuration member nodes of each cluster sends the collected information to the cluster headers. Cluster headers combine high quality information and data by a merge operation, thereby saving energy by reducing the number of packets sent to the base station [7].

Sensor nodes in the LEACH protocol consume energy in various ways, such as by data transmission, reception, cluster configuration, and data merging/compression, these processes, consume high energy for data transmission and reception. The following formula is used for data transfer.

$$E_{Tx}(k,d) = E_{Tx-elec}(k) + E_{Tx-amp}(k,d) = \begin{cases} k \times E_{elec} + k \times \in_{fs} \times d^2, & d < d_0 \\ k \times E_{elec} + k \times \in_{mp} \times d^4, & d \geq d_0 \end{cases}$$

The difference in energy consumed is due to bit size (k), distance (d) and threshold (d_0). Therefore, reducing the bit size of data transmission/reception while using energy efficiently is necessary [6].

LEACH member nodes transmit data without further processing, thereby resulting in high energy consumption.

3 Differential Data Processing

3.1 Differential Processing in Member Nodes

The means to handle the difference of member nodes is to first configure the cluster and reduce the size of the data transferred by using the data collected previously and the differential processing of the currently collected data. The member node conducts differential processing after data collection and sends it to the cluster header. After initial data collection, the member node passes the collected data along with the node ID to the cluster header without any difference processing and then stores the collected data in the repository. The next communication handles the difference from the value in the store and sends it to the cluster header along with the node ID. The member node updates the repository value with the last collected value.

In Fig. 1(a) shows the difference processing method of the data collected five times, and Fig. 1(b) shows the data transmitted in the cluster header. As can be seen the initial data 104 transmit the entire data in the cluster header. Then, the data collected from the communication only carries the difference value from the previously collected data. The difference values are 1, 2, 0 and −2. Thus, energy efficiency is improved by reducing the bit size of the transmitted data except for the initial data transmission.

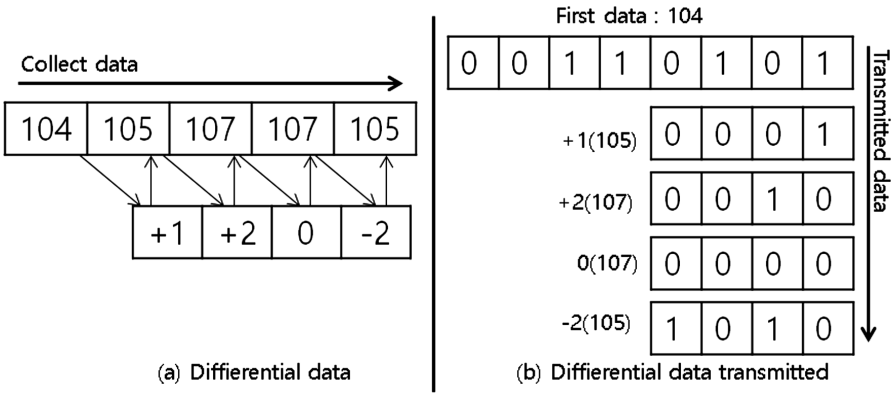


Fig. 1. Processing data

3.2 Aggregation Processing of Cluster Headers

The cluster header aggregates the data received by the member node and sends it to the base station. The proposed aggregation process uses the minimum value of the data received from the member node as reference data. When various data are differentially processed with the reference data, data other than the reference data excludes negative signs, such that the size of at least one bit is reduced. The smallest reference data are allocated and transmitted on the first packet.

The data structure transmitted from the cluster header to the base station is shown in Fig. 2. In differential data processing, the packet size of the first value is at least one bit larger than the value of the other value.

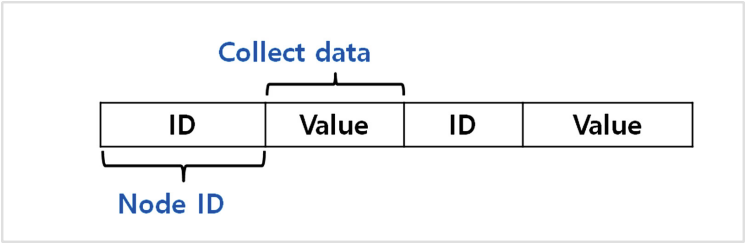


Fig. 2. Data structure transmitted

When the node ID value is −1 when transmitting from the cluster header to the base station, a special ID value indicating that the state of the member node is not communicable is transmitted. In this case, the value is the node ID.

Table 1 shows the three sets of data sent from the cluster header to the base station based on the data provided every minute at the weather station [9]. If the size of the data received from the member node in the proposed set is 0, it is not transmitted to the base station.

Table 1. Data sample

Minute	LEACH
01	{1:166,2:71,3:104,4:142,5:132,6:87,7:92,8:107,9:112,10:130,11:132,12:91,13:32, 14:17,15:138,16:15,17:75,18:81,19:46,20:60}
02	{1:166,2:71,3:103,4:142,5:132,6:88,7:92,8:106,9:113,10:130,11:132,12:89,13:31, 14:17,15:139,16:15,17:75,18:81,19:47,20:60}
03	{1:167,2:71,3:102,4:142,5:132,6:89,7:93,8:105,9:113,10:129,11:129,12:90,13:31, 14:17,15:139,16:15,17:75,18:81,19:47,20:60}
Minute	Differential data processing
01	{1:166,2:71,3:104,4:142,5:132,6:87,7:92,8:107,9:112,10:130,11:132,12:91,13:32, 14:17,15:138,16:15,17:75,18:81,19:46,20:60}
02	{12:-2,3:1,8:1,13:1,6:3,9:3,15:3,19:3}
03	{11:-3,3:2,8:2,10:2,1:4,6:4,7:4,12:4}

4 Performance Evaluation

4.1 Experiment Environment

In the simulation environment, 200 sensor nodes were randomly arranged in a 300 × 300 m^2 field using a MATLAB program. The base station is located on the X-axis (150 m) and on the Y-axis (50 m). The energy parameters of the set LEACH protocol are shown in Table 2. Furthermore, the data used in the experiment are the temperature data every minute collected from November 1, 2017, as provided by the Korea Meteorological Administration [9].

Table 2. LEACH energy parameters

Name	Parameter
Initial energy of node	0.5 J
Transmit amplifier (d < d0)	10 pJ/bit/m^2
Transmit amplifier (d > d0)	0.0013 pJ/bit/m^4
Transmitter/Receiver	50 nJ/bit
Data aggregation	5 nJ/bit

4.2 Performance Evaluation

Figure 3 shows the results of LEACH and differential data processing. As a result, the size of data transmitted from the cluster header to the base station is 360 bits in LEACH and 113 bits in DDP. Therefore, the average size of the data transmitted to the base station is reduced by 30% compared to LEACH.

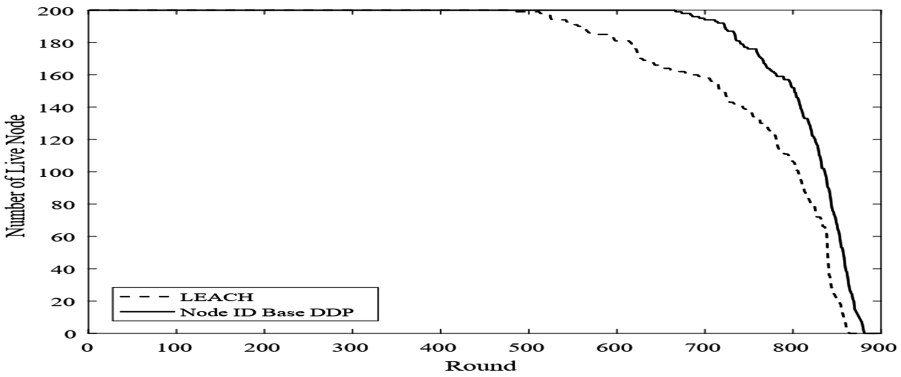


Fig. 3. LEACH vs DDP

Table 3 shows the number of rounds per node that died in the experiment. When the first node dies, LEACH is 483 times and DDP is 666 times. The difference between LEACH and DDP for the first node is 183 rounds. When the 100th node dies, DDP is 30 rounds later than LEACH. When the last node dies, DDP needs 17 more rounds than LEACH. Hence, experimental results show that DDP is more energy efficient than LEACH.

Table 3. Number of dead rounds of the node

Kinds	Dead node		
	First node	100^{th} node	Last node
LEACH	483	805	863
DDP	666	835	880

5 Conclusion

In this paper, we propose a differential data processing method for energy efficiency in wireless sensor networks. In the proposed method, the difference between the previous collected data and the currently collected data at the member nodes of the cluster is transmitted. We propose a merging method that performs differential processing based on the minimum data in the cluster header.

Experimental results show that on average, the data size of the sensor node is reduced by 30% via the DDP.

References

1. Kumarawadu, P., Dechene, D.J., Luccini, M., Sauer, A.: Algorithms for node clustering in wireless sensor networks: a survey. In: 4th International Conference on ICIAFS 2008, pp. 295–300. IEEE (2008)
2. Heinzelman, W.R., Chanrakasan, A., Balakrishnan, H.: Energy-efficient communication protocol for wireless microsensor networks. In: Proceedings of the 33rd Hawaii International Conference on System Sciences, Maui, USA, pp. 1–10, January 2000
3. Deosarkar, B.P., Yadav, N.S., Yadav, R.P.: Clusterhead selection in clustering algorithms for wireless sensor networks: a survey. In: International Conference on ICCCN 2008, pp. 1–8. IEEE (2008)
4. Heinzelman, W.R., Chandrakasan, A., Balakrishnan, H.: Energy-efficient communication protocol for wireless microsensor networks. In: Proceedings of the 33rd Annual Hawaii International Conference on System Sciences. IEEE (2000)
5. Sadler, J.C.M., Martonosi, M.: Data compression algorithms for energy-constrained devices in delay tolerant networks. In: SenSys, pp. 265–278 (2006)
6. Akyildiz, I., Su, W., Sankarasubramaniam, Y., Cayirci, E.: Wireless sensor networks: a survey. J. Comput. Netw. **38**(4), 393–422 (2002)
7. Rajagopalan, R., Varshney, K.: Data-aggregation techniques in sensor networks: a survey. IEEE Commun. **8**(4), 48–63 (2006)
8. Shin, D.H., Kim, C.: Data compression method for reducing sensor data loss and error in wireless sensor networks. J. Korea Multimedia Soc. **19**(2), 360–374 (2016)
9. Weather data public portal. Data.kma.go.kr (2018). https://data.kma.go.kr/. Accessed 08 Apr 2018

Performance Improvement on Object Detection for the Specific Domain Object Detecting

Hyungi Hong and Mokdong Chung(✉)

Department of Computer Engineering, Pukyong National University, Busan, Republic of Korea
heangi92@pukyong.ac.kr, mdchung@pknu.ac.kr

Abstract. Currently, various studies are underway to improve the performance of Object Detection technology. In the proposed system, rather than classifying all fields as a whole, we want to check the performance of the domain by performing additional learning on the specific domain. In this paper, we propose the performance improvement in the object detection system for detecting human in the coastal image that shot by drone.

Keywords: Object detection · Single Shot MultiBox Detector · Human detection · Computer vision · Image classification

1 Introduction

Recently, various and active researches for real-time object detection are being carried out. These studies suggest various ways to improve near real-time speed and accuracy. Most of the major studies currently aim to classify a various category of objects in a comprehensive range. To do this, Common Object in context (COCO) and Pascal Visual Object Classes (Pascal VOC) provide images containing objects of various categories. Using this image to learn each object detection techniques weight value and measuring the weight Mean Average Precision (mAP) for comparing performance.

Accuracy and speed are both important factors for systems that reliably detect real-time specific objects, such as on-site object detection for emergency rescue. In this case, recognition failure problem is more serious than the misrecognition of objects in the image. Therefore, it is important to decrease the recognition failure rate by learning objects through various types of learning data. In this paper, we aim to classify people. When the various postures and only part of the body are in the image, the classification rate of people drops. Especially in the case of the coast, there is possibility that a person enters the water and only a part of the human body is photographed and the recognition rate is lowered and not recognized. In order to solve this problem, we try to lower misrecognized rate and recognition failure rate in classification by labeling only part of body in labeling work.

In this paper, the drone takes a video of the ground and the coast in various environments and train through the additional image to classify human in the image with higher accuracy. We aim to train two types of weight value. We just use new

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 153–158, 2020.
https://doi.org/10.1007/978-981-13-9341-9_26

collecting dataset for training among them. Another one is trained by existing weight and collecting images. Comparing and evaluating existing weight and these two new trained weights for higher accuracy object detection.

The construction of this paper is as follows. Section 2 introduces the related work, Sect. 3 describes system design. In Sects. 4 and 5, we implement the system and evaluate it. Finally, in Sect. 6 shows the conclusions and future work.

2 Related Work

2.1 Single Shot MultiBox Detector

Single Shot MultiBox Detector (SSD) is kind of Object Detection technology. One of SSD variations, SSD512's test result of Pascal Voc 2007, mAP is 76.8%, speed performed 22 FPS [1]. In addition to You Only Look Once (YOLO) [2], it is mainly used as a real-time object detection system by ensuring higher speed than the existing Faster R-CNN.

In this paper, we used the existing weight trained by COCO 2017 dataset. And, we used SSD for training using new collecting dataset and evaluating the result.

2.2 Object Detection Related Research

Performance enhancement research for real-time object detection technology are being actively carried out through various directions. Technically, SSD [1] and YOLO [2] mentioned above have greatly improved speed through changes in algorithms and system structures. In addition to this, there have been various attempts to improve performance by optimizing recognition cell size, error optimization and elimination of duplicated calculations [3], learning through YOLO and confirming results with various devices [4].

3 System Design

In this paper, we aim to develop a system to more accurately detect people in images captured by a drone in coast. To do this, we collect a lot of coastal shots by drone and make dataset by labeling people in images. The collected datasets are trained through SSD and made weight value. In addition, the weight of the newly collected data set is added to the learned weight by the existing COCO 2017 dataset, and the performance is compared.

Fig. 1. Object detection server system design

Figure 1 is design of Object Detection server system. This server has two purposes that collect real drone shot image and shoot real environment. Server consists of three parts. First, Interface part communicates with the drone Ground Control System (GCS) to transfer and store images. Second, Object Detection Module detects specific domain in image using SSD and trained weight. Last, Streaming part stream result video to web page for check result anywhere. Drone can take various environment photo like closeup and high altitude.

4 Implementation

Development environment of this system is as follows. The Operating System used Linux Ubuntu 18.04 and vehicle language is python 3.6. For object detection, we used Tensorflow and Object Detection API. We choose SSD inception v2 in Object Detection API for high speed. Hardware spec is intel® i7-6700 CPU and GTX-1080 GPU.

4.1 Data Collection

In the case of detecting people in coast image taken picture by drone in sky, we can confirm the low recognition accuracy using existing weight trained by COCO 2017 dataset. In COCO 2017 dataset, only the images were composed that large in size and clearly identified with the naked eye. Therefore existing weight hard to classify small and various shape people. In order to solve this problem, we collected images of a person who seemed to be small from a distance and images of a body part of a person swimming in the coast. A total eight videos were collected from Google and Youtube. We divided these videos into 48 images. And we labeled 387 human images from this image and made xml file. We converted xml to csv, and csv to record file that our training dataset.

4.2 Training

We trained weight based on SSD and classify only one category that people for high performance and easy evaluation. We trained total two weights, one of both using only collected dataset record file. Another weight using existing COCO 2017 dataset weight trained by SSD inception V2 and collected dataset. During training, we confirmed low loss value and stopped training. We trained weight 20,000 steps. Finally, we convert checkpoint file to weight file.

5 Evaluation

System evaluation was confirmed by object detection result of images. We compared existing COCO 2017 dataset weight trained by SSD, newly weight trained by collecting dataset, additional training by newly collecting dataset on existing weight.

In this section, we evaluate performance using factors of precision and recall. And we hope to check object detection performance in specific domain, and we use original video that extracts images.

Fig. 2. Object detection result using COCO 2017 dataset weight and additional training weight

Figure 2 is the result of object detection in high altitude using existing COCO 2017 dataset weight and additional training weight. We could confirm some box and labeling results. Left figure had so low accuracy. If we applied a threshold of more than 30%, Then we checked just six box and labeling detection result. But, right figure that result of additional training weight confirmed a lot of box and labeling.

Table 1. Comparison of each weight precision and recall in high altitude situation

	SSD COCO2017	Collecting dataset	Additional training
Precision	100%	94%	95%
Recall	12%	44%	**46%**

Table 1 is evaluation result in high altitude. Original weight detected only one person. This weight value trained by people-centered photo. Another two weights detected nearby 40%. Most of remaining people located behind shadow and water. Thus it is hard to detect edge. As a result, additional training weight had good performance in the high altitude.

Fig. 3. Closeup image detecting result

Figure 3 is each weight's closeup image detecting result. It was original weight, collecting data weight, additional training weight in order. COCO 2017 dataset weight easily detected human in closeup image. We confirmed performance of collecting dataset weight, it has very low accuracy. This dataset did not contain closeup images. We predicted low accuracy in human-centered image. Additional training weight also has low accuracy. Additional training too many affected original weight values.

Table 2. Comparison of each weight precision and recall in closeup situation

	SSD COCO 2017	Collecting dataset	Additional training
Precision	83%	100%	100%
Recall	**40%**	16%	16%

Table 2 is evaluation result in closeup situation. In this situation, Original weight easily detected big things. oppositely, Another weights hard to detect targets.

In this section, we want to compare three weights value quantitatively. Therefore we count number of people and box in image. And also check each box state that true positives, false positives, false negatives, and true negatives. And then we calculate Precision, Recall and compare it.

As a result, we confirmed good performance of additional training weight in high altitude situation. But it has low accuracy in closeup situation because of affected by too many additional training. And also, it is hard to decide choosing training steps. Too different shape of domain interrupts each other weight. Thus, we proposed use two weights that is original one and collecting one for human detection.

6 Conclusion

In this paper, we improved performance of object detection server using additional weight. Object taken at different angles were difficult to recognize like another angle. Therefore we trained new weight value using object image taken at different angles. As a result, we increased recall value. It meant that we collect each situation's weight values and labels, then we can detect object in various angles.

In the further work, we are training various weight values for each labeling situation. And then, we hope to make high accuracy object detection system using combined weight value.

Acknowledgment. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF2017R1D1A1B03030033).

References

1. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, D., Berg, A.C.: SSD: single shot multibox detector. In: Computer Vision-ECCV 2016. LNCS, vol. 9905, pp. 21–37 (2016)
2. Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. arXiv:1612.08242 (2016)
3. Choi, H., Kim, S., Bae, G., Noh, Y., Lee, D.: Real-time object recognition and reporting system. In: Korea Software Congress 2017, pp. 2107–2109 (2017)
4. Lee, J., Kim, D., Kim, J.: CPU-based YOLO performance enhancement technique for CCTV real-time object detection. In: Korea Computer Congress 2018, pp. 870–872 (2018)

Implementation of Automatic Adjustment of Font Size System on Smartphone

Kyoung Soon Hong[1], JiSu Park[2], Kang Hyun Kim[1], and Jin Gon Shon[1(✉)]

[1] Department of E-Learning, Graduate School, Korea National Open University, Seoul, South Korea
soonkh@hanmail.net, {hgrikim, jgshon}@knou.ac.kr
[2] Division of General Studies, Kyonggi University, Suwon, South Korea
bluejisu@kyonggi.ac.kr

Abstract. Adjusting the font size is a function of the smartphone for improving readability. This smartphone is easier to use than other function but is less useful due to the psychological factors experienced by users, such as annoyance or fear of malfunction. Adjusting the font size is difficult when you are on the move or cannot use both hands to hold the smartphone. Therefore, we need a system that can automatically adjust the font size under any circumstances to improve readability. In this paper, we implement a system, in which the font size is automatically adjusted according to the distance between the user and the smartphone, without additional device operations. By automatically adjusting the font size, this system can increase readability while reducing difficulties in operation.

1 Introduction

Study to improve the readability of smartphones in devices and their various functional have been published. Among them, the font size can be adjusted by a simple touch or button operation. However, adjusting the font size by repeatedly adjusting the setting each time may prove to be difficult, especially when the user is on the move. People with disabilities and older people with reduced motor abilities can also feel more discomfort [1]. Therefore, to make it easier for users to control the font size, the manipulation of smartphones can be minimized, thus allowing the device to react according to the user's environment.

The automatic adjustment of font size (AAFS) system design, which is proposed in this paper, adjusts the font size according to a determined distance. Therefore, the user can adjust the size of the font that is easy to read by simply changing the distance of the smartphone from one's body, without needing to manipulate it anyway.

2 Factors Affecting Readability

The factors affecting readability are classified into visual, environmental and personal. Visual factors are the font characteristics seen by the user, such as font type, font size, font spacing, and sentence spacing. Environmental factors include the size of the

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 159–164, 2020.
https://doi.org/10.1007/978-981-13-9341-9_27

smartphone and the brightness of the user's surroundings. Personal factors are the user's characteristics, such as physical state, age, or disability. Table 1 shows the summarizes the three factors that degrade the reading environment.

Table 1. Summarizes the three factors that degrade the reading environment.

Readability inhibition factor	Contents
Visual factor	– Font type – Font spacing, between the lines, length of the line
Environmental factor	– Type of device to read – Lighting of reading space
Personal factor	– Individual vision difference – A degree of physical aging – A degree of eye disease

Among the three factors that hinder readability, the environmental and visual factors vary depending on where the smartphone is being used. Of all Internet users, 96.6% use smartphones or smart pads, and this shows that there are various places where smart devices can be used [2]. In addition, visual acuity decreases by 20% when moving, compared to when there is no movement [3]. When the reading environment is degraded due to environmental factors, it is possible to improve readability by merely adjusting the font size. However, using this feature is only 40% effective, which indicates that the operation of the device should be made easier.

Even if a user cannot use both hands or has a disability, smartphone operations should still be easy. In order to improve the convenience of using a smartphone while moving, the limitations of smartphone one-hand operation should be studied and analyzed [4].

Presbyopia, disease, and disability are personal factors that closely affects one's reading environment. A user adjusts the font size and brightness of the device to solve the personal factor. However, as the user's reading situation frequently changes, it may be troublesome to manipulate functions depending on each situation. Setting the font size to be larger than the normal size also reduces readability because it increases visual span. Visual span refers to the range of characters that can be accurately recognized at once without eye movement [5]. Frequent scrolling also negatively effects readability, so proper font size is required [6]. The user sets the appropriate font size for him/herself. Therefore, setting based on personal factors is not done frequently, and font size is often adjusted by environmental factors during movement.

3 Implementation of AAFS System

3.1 Distance Between User's Face and Smartphone

The core function of the AAFS system is automatic font size adjustment based on the distance between the user's face and the smartphone, without additional user

manipulation. The AAFS system uses the camera module embedded in the smartphone to detect the face of the user and extracts the number of pixels of the face length. The measured face length (L_{MF}) differs depending on the distance between the user's face and the smartphone's camera lens. Figure 1 presents the schematic diagram for calculating the distance between the user's face and the smartphone.

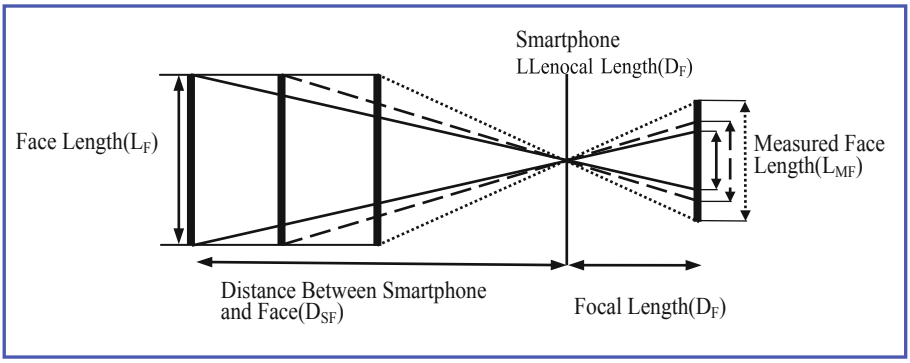


Fig. 1. Schematic diagram of distance measurement between user's face and smartphone

The L_{MF} determines the number of pixels in the face image displayed on the smartphone screen. When the user looks at the screen to read the text, the front camera is activated and the smartphone detects the face and obtains the number of vertical pixels in the detected face. The L_{MF} is the smartphone screen size divided by the number of pixels in the device, multiplied by the number of pixels in the smartphone face length. The L_{MF} obtained in Eq. 1 is the length value (in mm) of the face displayed on the smartphone screen.

$$L_{MF} = \text{Number of Face Length Pixel} \frac{\text{Screen Length}}{\text{Number of Device Pixel}} \tag{1}$$

The focal length (D_F) is a fixed distance according to the size of the image sensor in the smartphone camera module, but the image sensor is not used in the AAFS system. Therefore, the focal length is recalculated to the size of the smartphone screen. The calculation of the D_F is obtained by the ratio of the image sensor and the smartphone screen size. Equation 2 is the formula for obtaining the D_F.

$$D_F = \text{Focal Length of Device} \frac{\text{Device Screen Length}}{\text{Sensor Length}} \tag{2}$$

Meanwhile, the distance between the smartphone and the face (D_{SF}) can be obtained by multiplying the face length (L_F) by the D_F and dividing it by the L_{MF}, as shown in Eq. 3.

$$D_{SF} = \frac{L_F D_F}{L_{MF}} \tag{3}$$

Here,

D_{SF}: Distance between smartphone and face
D_F: Recalculated focal length
L_F: Face length
L_{MF}: Measured face length.

3.2 Automatically Adjusted Font Size According to Distance

The font size is automatically adjusted according to the distance. The size of the letters according to distance is based on the nearest point that can be seen by a user with presbyopia. The change interval of the automatically adjusted font size is from 220 mm to 310 mm in the first step and from 320 mm to 410 mm in the second step.

In the third step, the automatically adjusted for size is 420 mm and larger. As the length of the arm is limited, the third step is the maximum distance. Table 2 presents the distance value and the character size according to step, and it is the variable name that defines each value. Figure 2 presents the flowchart of the automatic font size control based on Table 2.

Table 2. Parameters of simulations.

Step	Distance	Distance variable	Font size	Font size variable
Step 0	−210 mm	D_0	12Pt	F_0
Step 1	220–310 mm	D_1	13Pt	F_1
Step 2	320–410 mm	D_2	14Pt	F_2
Step 3	420- mm	D_3	15Pt	F_3

4 Result of Implementation

The number of pixel of face length is obtained by extracting the facial area from the smartphone screen. The image size of face in smartphone depends on whether the user's face is near or far from the smartphone. It then calculates the distance value between the smartphone and the face and automatically adjusts the font size according to the distance value. The distance value between the smartphone and the face is calculated using Eq. 3. Figure 3 shows a screen that implements a system that is automatically adjusted according to the distance between the smartphone and the face.

In the implementation of system screen, the detected face is displayed as a rectangle, and the number of pixels of the face is expressed by face height. Steps 0, 1, 2, and 3 indicate the distance steps, and the letters at the top of the screen show the changes in the font size in steps.

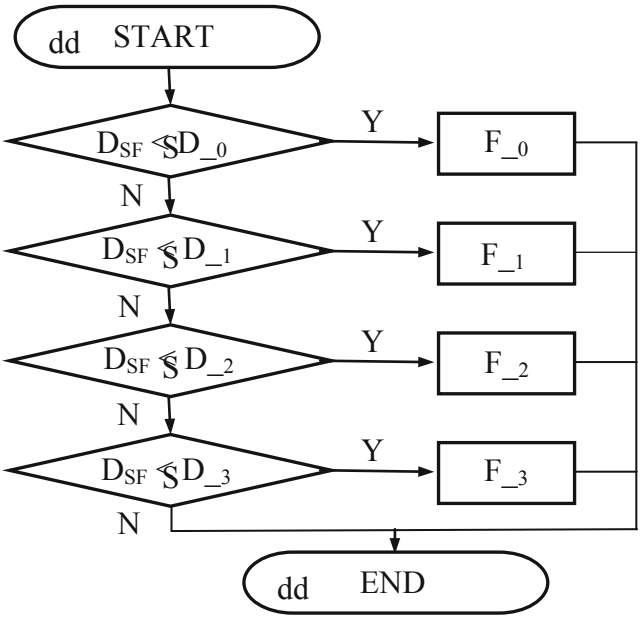


Fig. 2. Flowchart of automatically adjusted font size according to distance

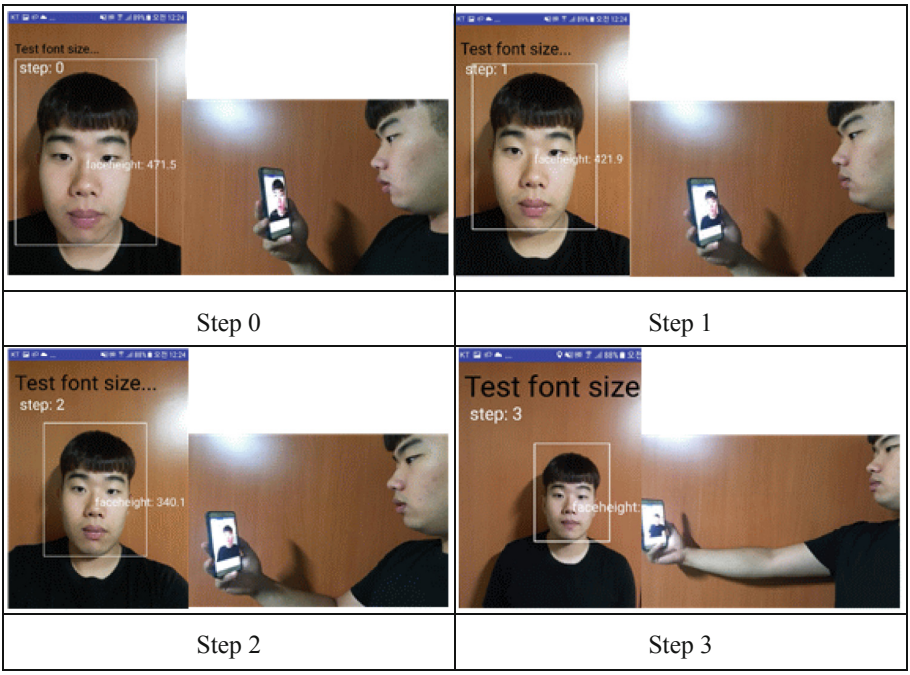


Fig. 3. Snapshot of AAFS system

5 Conclusion

The AAFS system is designed so that the font size is automatically adjusted without user manipulation, here, the font size can be changed by the distance between the user's face and the smartphone. In this method, the distance of the smartphone to the user's face is obtained using the built-in camera, and the font size is automatically adjusted according to the distance value.

Given that the AAFS system adjusts the font size in any environment adjusting the font size and returning it to the original size becomes easier. Depending on the situation, this method can increase the usability of smartphones, especially for those who have difficulties in performing one-hand operations or among the elderly who have difficulties in handling devices.

References

1. Diaz-Bossini, J.-M., Moreno, L.: Accessibility to mobile interfaces for older people. Proc. Comput. Sci. **27**, 57–66 (2014)
2. Survey on the Internet Usage. Korea Internet & Security Agency (2017)
3. Conradi, J., Busch, O., Alexander, T.: Optimal touch button size for the use of mobile devices while walking. Proc. Manuf. **3**, 387–394 (2015)
4. Oh, S.H.: Usability evaluation factors of smartphone one-handed operation interface. Thesis Paper, Hongik University (2014)
5. Jeong, B.: Measurement of visual span and reading abilities in korean reading. Thesis Paper, Chung-Ang University (2015)
6. Sanchez, C.A., Goolsbee, J.Z.: Character size and reading to remember from small displays. Comput. Educ. **55**(3), 1056–1062 (2010)

Development of Usable Communication Reasoning System Using Connection Information for Efficient Conversion of Connection

Taehoon Koh[1], Yonghoon Kim[1], Kamyoung Park[2], and Kyungryong Seo[1,3(✉)]

[1] Department of Marine Lab., SUNCOM Co., #602, NIFC, 111, Hyoyel-ro, Buk-gu, Busan, Republic of Korea
thkoh@suncom.co.kr, kimyhjava@pukyong.ac.kr
[2] Korean Register, 36, Myeongji ocean city 9-ro, Gangseo-gu, Busan, Republic of Korea
Kamyoung@krs.co.kr
[3] Department of Computer Engineering, Pukyong National University, 45, Yongso-ro, Nam-gu, Busan, Republic of Korea
krseo@pknu.ac.kr

Abstract. In the waters where various communications of Wi-Fi, VHF, LTE, and Satellite should be used, economical and performance aspects are considered. Offshore vessels use a smaller number of communications, but there are many difficulties in supporting communications that meet customer requirements. In particular, in the case of a ship operating a long distance, the burden of the communication cost may affect the overall transportation cost, and therefore, a significant cost should be incurred if proper communication conversion is not achieved. In this paper, we describe the problem of applying the existing communication connection policy and discuss the development of a system that predicts and reflects the problem using the communication collection information in the ocean.

Keywords: Communication · Geo-fencing · Random variable · e-Navigation

1 Introduction

LTE, Wi-Fi and Wibro are the most common types of terrestrial communications in 2011, with 4G [1] being the first year [2]. In particular, LTE provides all services including IP-based data using TCP/IP. It can be provided faster [3]. Based on these technologies, the Ministry of Public Administration and Security (MSTF) in 2011 introduced the optimal communication technology method called PS-LTE (Public Safety LTE) in Korea through the feedback of disaster network demand organizations [4] (IRIS) system including LTE-R (LTE-Railway) in the existing fixed-line network and VHF system, and plans to integrate wired/wireless by 2022 and unify the whole country in 2016 [5]. In addition to this, in order to reduce marine accidents caused by

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 165–171, 2020.
https://doi.org/10.1007/978-981-13-9341-9_28

human factors of IMO (International Maritime Organization), we introduced the e-navigation system and it was proposed to develop for the first time in 2005. After 8 years of discussions (NAVIGATION PROJECT) and started standardization and development from October 2016 to 2020 [6].

However, wireless communications at sea have environmental limitations. The difficulty of installing fixed structures and the problem of no transmitting and receiving scatterers [7] have selective restrictions on communication equipment depending on the distance between the land and the ship. To the extent required by law. The problem of the communication range can be solved somewhat by using the satellite, but it has a problem in connection with the wireless communication method developed around the base station and a problem about the communication cost. In order to satisfy these various conditions, various communications equipment such as VDES (VHF), Wi-Fi, HF/MF, 3G/LTE and Satellite are used.

Such an environment requires efficiency and reliability for connection of various communication equipment and requires an integrated marine communication gateway for reliability of connection. In addition, although the requirements for users with different characteristics such as communication service and economic cost are increasing, it is a reality that there is a shortage in relation to the application of the conversion algorithm which satisfies this sufficiently. Commercial models developed around LTE and 3G convergence [8–11] interfere with additional connections due to difficulties with non-TCP/IP connections, In most cases, users are required to perform a direct conversion and cannot provide information about each communication. Therefore, the user must review and change the appropriateness and economic situation to each communication connection. Therefore, in this paper, to develop a system for predicting connection communication without examining communication connection in existing communication switching policy, based on communication information according to location, communication intensity for each communication is calculated as a random variable and applied. Section 2 describes the related research, and then Sect. 3 provide the implementation and evaluation of the system, and Sect. 4 will discuss the conclusion.

2 Related Work

2.1 Expectation of Random Variable

The expected value of a random variable can be estimated as an average value. The overall distribution of X is generally too difficult to express this information. This is because the distribution of the random variable X contains all probabilistic information about X. The average or expected value can be useful for providing information about where X is expected [12].

Normally, the expectation of X, denoted by E(X), is a number defined as follows (1). Expectation for a continuous distribution is (2).

$$\mathrm{E}(\mathrm{X}) = \sum\nolimits_{All\ x} xf(x) \tag{1}$$

$$\mathrm{E}(\mathrm{X}) = \int_{-\infty}^{\infty} xf(x)dx \tag{2}$$

2.2 Geo-Fencing

Geo-fencing is a function of software programs that use geographic bounding systems (GPS) to define geographic boundaries. This will usually warn you when you are within or outside your defined bounds.

Among many geo-fencing applications, Google Earth [13] is a typical example. Google Earth allows users to define boundaries above a satellite view for a specific geographic area. By defining these boundaries, we use this technique to determine whether the actual user is in the area or not.

3 Implementation and Evaluation

3.1 Implementation

In order to learn, there are a large number of algorithms according to each learning method, and guidance learning is conducted in data processing for machine learning. In learning maps, it is important to process the data to be learned. General learning variables can set communication intensity and coordinates for each type of communication, but it can be a problem at sea when setting judgment tag to be applied. In other words, commonly used coordinates are divided into Geodetic Longitude and Latitude and Projected Coordinate, and the general method used in the ocean is the longitudinal latitude coordinate system. In addition, the notation of the latitude and longitude index system is classified into [degree] notation and [degrees/minute] notation. In [figure] notation, Degrees/min], there is a difference of 2.5 m in distance to 0.1 m in latitude 0.1″ and 0.1 m in latitude. The problem is that in the collection of training data, the number of decimal places or [degrees/minutes] to the number of seconds is used as the range in the notation. If you measure the communication strength by measuring 0.00001 to the decimal point in the actual sea, you will have to collect huge amount of data. If so, reducing the number of decimal places can cause problems for each communication decision over a wide range. In other words, when the LTE, Satellite, and VHF communication are mixed at a certain point, the communication intensity also occurs at various values, and it is difficult to decide which communication is reasonable to select. In addition, there are some cases where a ship is operated only for a given route, and it is difficult to estimate the decision value for the area other than the route.

Therefore, it can be said that probability is used when various communication and communication intensity are derived at a certain point, but it is said that it is the same as the case where the front and back of the coin come out at 5:5 and may be more difficult in an unstable communication environment.

Therefore, we processed the data to be learned by assigning the appropriate tag value using the random variable for each communication.

The signal strength of each communication for each coordinate is iteratively extracted. When the probability variables L, V, S, H for LTE, VHF, Satellite and HF communication have a probability distribution, $\mathrm{P}(l), \mathrm{P}(\mathrm{v}), \mathrm{P}(\mathrm{s})$ *and* $\mathrm{P}(\mathrm{h})$ expected value is (3).

$$E(\mathrm{l}) = \sum_{All\ x} lP(l),\ E(\mathrm{v}) = \sum_{All\ x} vP(v), E(\mathrm{s}) = \sum_{All\ x} sP(s), E(\mathrm{h}) = \sum_{All\ x} hP(h) \tag{3}$$

In the present study, we apply discrete random variables, but in case of continuous, (4) should be applied.

$$E(\mathrm{l}) = \int_{-\infty}^{\infty} lP(l), E(\mathrm{v}) = \int_{-\infty}^{\infty} vP(v), E(\mathrm{s}) = \int_{-\infty}^{\infty} sP(s), E(\mathrm{h}) = \int_{-\infty}^{\infty} hP(h) \tag{4}$$

However, the communication according to each coordinate at sea is highly related to the temperature and humidity airflow, and the values measured according to the weather changes most closely. Therefore, when a time series analysis is performed, the multivariate binding distribution for the calculated random variable should be considered.

If the n random variables (time series) $l_1, l_2, l_3, \ldots, l_n$ of the LTE communication have the discrete joint distribution $f_1(l_1, l_2, l_3, \ldots, l_n)$ and n variables are defined as m functions of $Y_1, Y_2, Y_3, \ldots, Y_m$, can be expressed as (5).

$$\begin{aligned} Y_1 &= r_1(l_1, l_2, l_3, \ldots, l_n) \\ Y_2 &= r_2(l_1, l_2, l_3, \ldots, l_n) \\ Y_3 &= r_3(l_1, l_2, l_3, \ldots, l_n) \\ &\cdots \\ Y_m &= r_m(l_1, l_2, l_3, \ldots, l_n) \end{aligned} \tag{5}$$

Then, for the set A in which $Y_1, Y_2, Y_3, \ldots, Y_m$ is defined for all $(l_1, l_2, l_3, \ldots, l_n)$ values, the joint probability function $\mathrm{g}(y_1, y_2, y_3, \ldots, y_n)$ of $Y_1, Y_2, Y_3, \ldots, Y_m$ can be calculated as follows (6).

$$\mathrm{g}(y_1, y_2, y_3, \ldots, y_n) = \sum_{(l_1, l_2, l_3, \ldots, l_n) \in A} f_1(l_1, l_2, l_3, \ldots, l_n) \tag{6}$$

Experiments were carried out to establish decision values for machine learning by using 1002 coordinates in total and setting communication intensity for each communication in each coordinate.

Figure 1 shows that the range of each communication is set and the simulation is conducted using it. The upper left-hand side represents Wi-Fi, the right side represents the range of satellite communication, the lower left represents LTE, and the lower right represents the range of VDES communication.

In the figure, R (Real) represents the measurement of the communication intensity, A (Average) is the average of the surrounding communication information, and E (Expected) is the calculation of expected value.

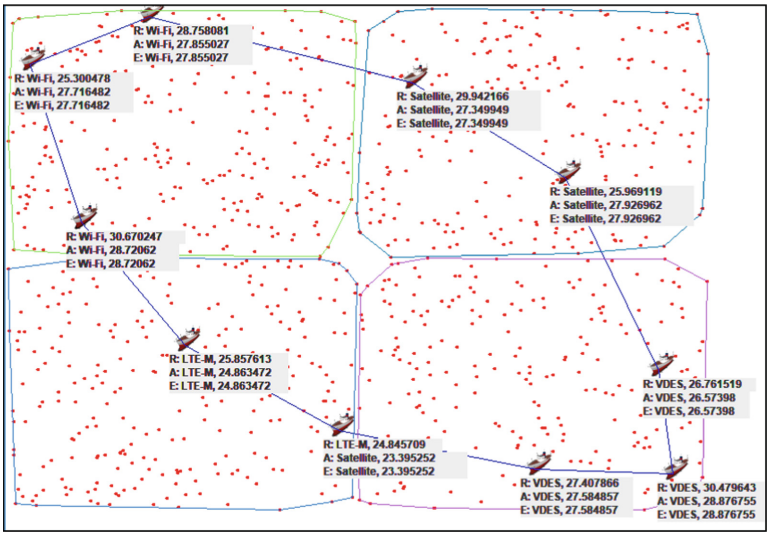


Fig. 1. Example of communication classification experimental

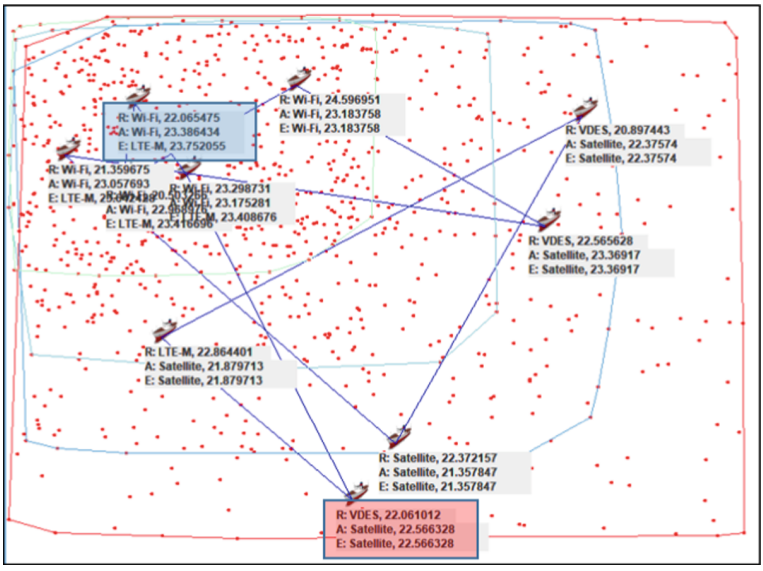


Fig. 2. Example of real-world simulation

In this paper, we can derive a communication decision value that is closest to R by analogy without using the existing R method, and it can be seen that both the A and E methods are approximate estimates. However, the actual sea area has a communication range as shown in Fig. 2, and the range of Wi-Fi, LTE, VDES, and Satellite is shown from the upper left. It can be seen that estimating A from the estimated value of the blue rectangle accurately estimates LTE while estimating LTE from E is assumed. In the red rectangle, the full range of satellites should be used. Is estimated more precisely.

3.2 Evaluation

In this study, we tried to estimate the communication within the range to obtain the decision value for machine learning. In this experiment, we use the general communication switching policy of marine, and the related matters are shown in Table 1 below.

Table 1. Experiment environment

Division	Contents	
Connection policy		Wi-Fi>Lte>VDES>Satellite
Communication switching time	Wi-Fi	3 s
	Wi-Fi → Lte	3 s
	Lte → VDES	3 s
	VDES → Satellite	4 s
Connection attempt		3 time
Effective communication strength		20 RSSI–25 RSSI

RSSI: Received Signal Strength Indicator
VDES: VHF Data Exchange System

In order to process the experiment, the results were measured with the calculation range and the random selection ship, and the accuracy with the existing method was confirmed. The results are shown in Table 1.

4 The References Section

Information gathering is also an important part of various communication decisions at sea, but analyzing the collected information is also an important part, and communication decisions near the land are difficult. Because it is possible to do all communication such as Wi-Fi, LTE, VDES, Satellite, etc., close to the land, if it goes through only connection review, it may take a lot of time for communication connection like experiment result, In the case of LTE, costs are incurred. In order to solve this problem, a simulator has been developed to develop a system for estimating using the collected communication information. In some cases, the result is different from the actual decision value. However, I think it's worth it.

In future work, we intend to develop a system that can increase the accuracy of estimation and reflect on machine learning and make a reasonable decision.

References

1. Advanced New Technology Information Analysis Research Group: Next-generation mobile communication LTE market trend and technology development strategy, JinHanMNB (2011)
2. K-ICT spectrumMap: Korea Broadcasting & Communications Agency, 2016 (2018). http://spectrummap.kr/stats/statsAnalysisView.do
3. Dahlman, E., Parkvall, S., Skold, J.: 4G: LTE/LTE-Advanced for Mobile Broadband, 2nd edn. Elsevier, Amsterdam (2014)
4. Park, H., Lee, S., Lee, K., Kim, D., Kwon, J.: National disaster safety network technology method. J. Korean Commun. Soc. Inf. Commun. **31**(10), 12–18 (2014)
5. Yoon, H., Park, J., Yun, Y., Seong, D.: Commercialization of LTE-R (railway radio communication network) and national railway network construction plan. J. Korean Commun. Soc. Inf. Commun. **33**(3), 66–73 (2016)
6. An, K.: A study on the improvement of maritime traffic management by introducing e-navigation. J. Korean Soc. Mar. Environ. Saf. **21**(2), 164–170 (2015)
7. Jang, J., Lee, S.R., Kim, D.H.: Spatial multiplexing using open-loop precoding in maritime communication environment with channel correlation and LOS. J. Korean Inst. Commun. Inf. Sci. **40**(7), 1397–1404 (2015)
8. Net Co., Ltd.: NCT, NET Co., Ltd., 2016 (2018). http://www.netjoiner.com
9. IDO-LINK Co., Ltd.: IDO-LINK Co., Ltd., IDO-LINK, 2018 (2018). http://www.ido-link.com
10. Locomarine: Locomarine YACHT ROUTER, 2016 (2018). http://www.yachtrouter.com
11. Digital Yacht Ltd.: Wireless Internet, Digital yacht, 2018 (2018). http://www.digitalyachtamerica.com
12. Morris, M.J.S., DeGroot, H.: Probability and Statistics, 4th edn. Pearson Education Inc., London (2012)
13. Rouse, M.: WhatIs.com, TechTarget, 1999 (2018). https://whatis.techtarget.com/definition/geofencing

Textual Classification to Distributional Representation Using Cohesion Devices

Yonghoon Kim[1], Sookhyun Jung[2], Jee-Won Hahn[2(✉)], and Mokdong Chung[1(✉)]

[1] Department of Computer Engineering, CS and AI Lab, Pukyong National University, 45, Yongso-ro, Nam-gu, Busan, Republic of Korea
kimyhjava@pukyong.ac.kr, mdchung@pknu.ac.kr

[2] Department of English Language and Literature, Pukyong National University, 45, Yongso-ro, Nam-gu, Busan, Republic of Korea
sookhyun31@gmail.com, jeewonh@pknu.ac.kr

Abstract. Research on the information retrieval model using neural networks has been actively progressed to the retrieval of information and the classification of documents. Various algorithms have been applied to realize this research, and weights have been implemented to words and documents. As information technology accelerates, we might be necessary to understand the exact meaning of documents through the advanced method of obtaining and analyzing the values. Discussions on the frequency of words directly related to these studies have focused on English learning in social sciences or on pure language. In this paper, we propose a method that can more accurately analyze the meaning of sentences by reasonably increasing the frequency of important words in sentences. The essential words are based on cohesion devices used to make context. Several tests were conducted to verify the performance and the effect of improving the semantic analysis system, and we confirmed its result information in the sea.

Keywords: Cohesion devices · Word2vec · Textual classification · Neural network · Reference

1 Introduction

Data analysis is divided into structured data and unstructured data according to the classification. It also can be separated into supervised learning and unsupervised learning through the learning method. Currently, unstructured data analysis using unsupervised learning is more actively performed. Unstructured data means data which meaning is not precisely defined with images, sounds and sentences. In particular, Information Retrieval [1] among textually unstructured data analysis is a significant factor in the information age. The mechanism of information retrieval reflects how often people explore a search word. Specifically, supervised learning in information retrieval refers to putting on a label to each document to match it with its proper label. Unsupervised learning is to learn and classify information corresponding to a specific query as features of a document. There is Latent Semantic Analysis (LSA) [2] as a

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 172–177, 2020.
https://doi.org/10.1007/978-981-13-9341-9_29

traditional analysis model and Latent Dirichlet Allocation (LDA) [3] using stochastic techniques. Recently, as the interest of neural network with Convolution Neural Network (CNN) [4] has increased, using this analysis is magnifying the information retrieval. The most dominant thing in the information retrieval is to classify the features of documents accurately. As for the similarity of words, numerous researchers have used various algorithms such as Continuous Bag-of-words (CBOW) employing N-gram and skip-gram [5], Word2vec putting the relation of surrounding on words, Paragraph2vec [6] applying a length of documents, and others. Therefore, the value of frequency is a critical part of this analysis because it is grounded in measuring the frequency of words and documents. However, the related discussion is still lacking. In this paper, we propose a semantic improvement of documents by replacing demonstrative nouns into specific words based on cohesion devices which contribute to the sentence structure and composition in English literature. Therefore, this analysis helps researchers to learn the characteristics of the document, and it will contribute to predicting similar documents and words. The order of this paper will be as follows. Section 2 describes the related research, and then Sect. 3 proposes the method of applying the cohesion device. Sections 4 and 5 provide the implementation and evaluation of the system, and Sect. 6 will discuss the conclusion.

2 Related Work

2.1 Cohesion Devices

Cohesion is linguistic features in the passage that contribute to its total unity to make it text. The linguistic connectedness occurs where the interpretation of some element in the discourse is dependent on that of another [7]. It indicates that the comprehension of a text is associated with perceiving relation of each element. Cohesion is achieved through linguistic devices such as reference, substitution, ellipsis, conjunction, and lexical cohesion [8]. Among them, this study concentrates on two elements; reference and repetition. Pronouns traditionally gain the topic-head position, which makes a contextualized flow of information. Lexical repetition is the principal means of explicitly marking cohesion in text [9]. Hence, we assume that replacing pronouns with nouns can increase cohesion.

Many previous studies have proved that a high-cohesion text improves participants' text-based comprehension. Manipulating degree of cohesion causes topic continuity, which results in participants better understanding texts [10]. Regardless of participants' level of reading proficiency, high-cohesion texts produce better scores in their reading performance tests [11]. These results report that a high-cohesion affects an individual's comprehension of a text. Therefore, this study hypothesizes that computers also can detect the topic of texts more accurately within a considerable degree of cohesion. It can be inferred that a high-cohesion can generate more precise results in computer classifying documents. These cohesion devices are divided by lexical cohesion devices and grammatical cohesion devices depending on vocabulary and grammar.

The lexical cohesion devices represent the previous information repeatedly and thereby enhance cohesion. On the other hand, the grammatical cohesion devices are to

point out and substitute the preceding information with deixis or to omit repeated words to display the relation of the previous information [12]. Representative types of cohesion devices are shown in Table 1 [7].

Table 1. Representative types of cohesion devices

Category	Device	Example
Reference	Pronoun	*Jeff slept until the alarm clock woke him up*
	Demonstrative	*Would you like this cake? I bought it this morning*
	Comparative	*This novel is lighter but more subtle*
	Definite article	*Last year we went to Devon for a holiday. The holiday we had there was the best we've ever had*
Substitution	Nominal	*Would you like this cake? I bought one this morning*
	Verbal	*You can't live on what they would pay you. You could do on twice as much, maybe*
	Clausal	*I didn't think she was exceptionally shy. She wasn't at one time, but she has become so recently*
Ellipsis	Nominal	*Which last longer, the curved rods or the straight rods? The straight (ones) are less likely to break*
	Verbal	*Have you been swimming? Yes, I have. (been swimming)*
	Clausal	*John is a coward Yes, but not I. (I am not a coward)*
Conjunction	Additive	*In addition, plus, moreover, furthermore*
	Adversative	*In contrast, on the other hand, however, nevertheless*
	Causal	*For this reason, thus, therefore*
	Temporal	*Then, afterwards, now*
Lexical	Repetition	*A boy is climbing the old elm. That elm isn't very safe*
	Synonym	*I've been to see my great aunt. The poor old girl is getting very forgetful these days*
	Collocation	*Why does this little boy wriggle all the time? Girls don't wriggle*
	General nouns	*Henry seems convinced there's money in dairy farming. I don't know what gave him that idea*

2.2 Tensorflow Softmax and Neural Network

Tensorflow is an open-source software library that performs numerical computations using data flow graph. The graph's Nodes represent numerical operations, and Edge represents multidimensional data tensors that move between nodes. It can use a CPU or GPU to run performance on a server or mobile device. It was developed primarily for machine learning and Deep Neural Network (DNN) research by Google Brain team, a Google AI research organization [13]. Softmax or Softmax regression is equivalent to replacing the output of the existing Neural Network with a probability distribution without using the sigmoid function. Neural Network of Artificial Neural Networks are

models for classification and prediction, and Multilayer Feedforward Networks is the most flexible method to grasp the complex relationship of response variables. Transfer functions include linear functions, exponential functions, and Logistic/Sigmoid functions. Sigmoid functions are used as typical examples [14, 15].

3 Method of Applying-Cohesion Devices

This study adopted cohesion devices to obtain a feature of each document. We collected data on Wikipedia about specific topics including Love, Hate, War, and Peace. We put them into five categories as Love, Hate, War, Peace, and War and Peace regarding that 'war' and 'peace' usually co-occur like a collocation. Next, the data is divided into two groups on the basis of existence of cohesion devices. As I already mentioned, as a reference indicates the previous information, it can be re-write as the preceding word or phrase.

4 Experiment on Document Identification

In the experiment on document classification, we should make the documents into two groups; with cohesion devices and without cohesion devices. First of all, we extracted a characteristic of each document and estimated a similar document through machine learning. To do this, we used to 27 documents out of 37 to train machine and utilized the rest of them, 10 documents, as verification data. Testing data was 4,085 words without cohesion device while it was 4,109 words with cohesion device. The tokens were 1,599 respectively. We confirmed that the result of accuracy based on machine learning was satisfactory in the previous research [16]. Therefore, we extracted similar words to pronouns and references using Word2vec, and then put weights on these words. The 240 Documents, 122,811 words, and 11,330 tokens were used in the experiment. The accuracy of the whole words was approximately 70% while the average accuracy without cohesion devices was 54% [17], which shows 16% of deviation. Subsequently, we examined 400 Documents, 202,093 words, and 15,425 tokens to increase the reliability of experiments. In order to achieve scientific credibility in actual experiments, 400 documents were arbitrarily mixed. 80% of them were used as training data, and 20% of them were verification data. Since the selection of both learning data and verification data will affect estimating accuracy, we should conduct experiments five times to control this variable.

5 Evaluation

In the earliest experiment of this study, we applied λ_d. The accuracy was 70% without the cohesion device. On the contrary, 90% accuracy was confirmed after applying the cohesion device. Among ten verification data composed of two documents in each topic, only one document in 'peace' was estimated in error. Thus, this study assumed that adopting cohesion devices made better result than without them through the first

experiment. Following this acceptable result, we employed word2vec extensively in the way we set a more weighted value on related words to cohesion devices. The result was remarkable. Thus, we experimented more data five times. To diminish the deviation of accuracy for fixing learning data and verification data, we performed experiments five times by randomly mixing and exchanging data. The results are in Table 2.

Table 2. The result between with cohesion and without cohesion

Division	240 document					400 document				
	#1	#2	#3	#4	#5	#1	#2	#3	#4	#5
With cohesion	64%	66%	73%	65%	63%	71.2%	78.7%	63.7%	71.2%	65%
Without cohesion	53%	52%	57%	56%	48%	70%	76.2%	66.2%	73.7%	62.5%

In some cases, the accuracy of without cohesion was high. In other words, imposing weight has an adverse effect on learning performance.

6 Conclusion

This study attempted to classify multiple documents based on machine learning. Since acquiring actual documents in machine learning showed the deficient result than we expected, we applied cohesion devices to see if the linguistic feature met the increased performance. Cohesion devices are a kind of linguistic tools to hold sentences together and give them meaning within context. It means almost every document has cohesion devices in its content. In addition, the study used word2vec. As the machine learning for classifying documents has usually relied on the frequency of words, insignificant words such as function words can record a high ranking. Even if these words do not represent the content of a document, machine learning can make the result based on the list of word frequency. This process may lead to the inadequate accuracy of distinction among documents. Thus, regarding this defect, we extracted words corresponding to cohesion device especially references using word2vec and replaced them into nouns. We presumed that using repetition of nouns might increase the weighted value of these words, which could improve learning performance. In the experiments, word2vec was not able to deduce precise references or pronouns from the sentences. However, it can detect some notable words that enhance the discrepancy of documents. These noteworthy words improved the weighted value. Meanwhile, the learning performance was also advanced. As a consequence, if we add algorithms which can find out cohesion devices more precisely, the learning performance might progress. In the further research, we will conduct research that can more accurately deduce the words corresponding to cohesion devices.

Acknowledgments. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF2017R1D1A1B03030033).

References

1. Bhaskar, M., Craswell, N.: Neural models for information retrieval, arXiv preprint arXiv: 1705.01509 (2017)
2. Deerwester, S.C., Dumais, S.T., Landauer, T.K., Furnas, G.W., Harshman, R.A.: Indexing by latent semantic analysis. JASIS **41**(6), 391–407 (1990)
3. Tyson, L.D.: Who's Bashing Whom?: Trade Conflict in High-Technology Industries. Peterson Institute (1993)
4. Razavian, A.S., Azizpour, H., Sullivan, J., Carlsson, S.: CNN features off-the-shelf: an astounding baseline for recognition. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE (2014)
5. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space, arXiv preprint arXiv:1301.3781 (2013)
6. Rong, X.: word2vec parameter learning explained, arXiv preprint arXiv:1411.2738 (2014)
7. Halliday, M.A.K., Hasan, R.: Cohesion in English. Longman, London (1976)
8. Birner, B.J.: Introduction to Pragmatics. Wiley-Blackwell, Oxford (2013)
9. Hoey, M.: Patterns of Lexis in Text. Oxford University Press, Oxford (1991)
10. Kim, S., Han, K., Cho, S.: The effect of cohesive devices on memory and understanding of scientific text. Korean J. Cogn. Sci. **13**(2), 1–13 (2002)
11. Lee, J.-S., Lee, J.-W.: Effects of text coherence on EFL learners' English reading performance. Engl. Lit. Lit. Teach. **20**(3), 321–339 (2014)
12. Ahn, S.: A study on cohesion apparatus usage patterns in Korean and English speakers. Saehan Engl. Lang. Lit. **59**(2), 163–186 (2017)
13. Google Inc.: https://www.tensorflow.org
14. Galit, S., Bruce, P.C., Stephens, M.L., Patel, N.R.: Data Mining for Business Analytics: Concepts, Techniques, and Applications with XLMiner. Wiley, Hoboken (2016)
15. Rumelhart, D.E., Hintont, G.E., Williams, R.J.: Learning representations by backpropagating errors. Nature **323**, 533–536 (1986)
16. Kim, Y., Hong, H., Chung, M.: Application of cohesion devices for improvement of distributional representation. In: Proceeding of the 14th International Conference on Multimedia Information Technology and Applications (MITA2018), 28–30 June 2018, pp. 84–87. Shanghai University of Engineering Science, China (2018)
17. Kim, Y., Chung, M.: Test on learning method for improving performance using cohesion devices. In: 2018 Proceeding of Korea Information Processing Society, vol. 25, no. 2, pp. 755–758 (2018)

Information Assurance Requirements for Software Controlled Measuring Instruments

As the Electric Energy Meter Case

Seung-hwan Ju[1(✉)], Kwang-jae Song[1], Sang-hoon Song[1], and Hee-suk Seo[2]

[1] Korea Testing Laboratory, 723, Haean-ro, Ansan, Republic of Korea
{seunghwan, song, shsong}@ktl.re.kr
[2] KoreaTech, 1600, Chungjeol-ro, Cheonan, Republic of Korea
histone@koreatech.ac.kr

Abstract. Legal metrology and instrumentation standard equipment is controlled by software. It is no wonder that the software is updated via the network. Accordingly, there are new problems such as a function stoppage and information leakage by a cyber-attack. In order to ensure that this software controlled instruments operate correctly, there is need to verify that the software is not tampered. To ensure that the software works correctly, there is need to provide the software requirements and appropriate verification methods. In this study, we have studied the construction of software authentication system for measuring instruments. We identified five goals for protecting and validating instrumentation software. We derive cyber-attack threats and scenarios, and propose software requirements to overcome them. It is expected that this research will be able to introduce not only electric power meter but also software certification of type approval measurement instruments.

Keywords: Fraud prevention · Instrument software · Legal metrology · Measuring instruments · Software certification · Software protection · Software requirement · Software validation · Software vulnerability · Energy meter

1 Introduction

Smart meter refers to an electronic energy meters that communicates with the service provider's network for purposes of monitoring and billing purposes. It mainly uses the amount of reactive power and power factor.

Which is more intelligent and various supplementary services than existing electronic energy meters. The most significant difference between a smart meter and a conventional mechanical energy meter is that it has real-time measurement and storage capabilities for a variety of data and that a power grid operating server is capable of real-time bidirectional communication with other intelligent devices in the household [1–3].

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 178–183, 2020.
https://doi.org/10.1007/978-981-13-9341-9_30

The Smart Meter is more intelligent and more versatile than today's conventional electronic watt-hour meters and communicates with the service provider's network. Smart meters are expected to contribute to solving all kinds of problems through Smart Grid by supporting the interactivity of Smart Grid in various aspects [4, 5].

The metrology system is a system that operates and maintains the accuracy and maintenance of the meter in the field where there is a conflicting measurement result by the meter. Furthermore, it aims at ensuring the fairness of measurement and protecting the lives of the people in areas where inaccurate measurement results may have negative effects on individuals or society. In this paper, we design software requirements and test scenarios for information assurance of the electricity meter.

2 Related Works

The smart metering is developing into a metering system service industry using the information and communication technology. Various technologies such as SW-based big data and cloud are applied and various application services are promoted in order to guarantee energy efficiency and consumer's choice of energy in order to prepare for the new system [6].

2.1 European Legal Meter System (MID)

In accordance with the EU blue guide in 1999, the EU has set up a legal metering system and evaluation system that prioritizes securing the interests of consumers in order to develop new markets in Europe. The EU Legislation Guidelines (MID, 2004/22/EC), it includes the related regulations, standards, and certification system for 10 types of statutory meters. The main content covered in the European legal metrology system deals with the requirements for software to ensure a fair commerce system and an evaluation system that can incorporate various technologies [7].

The EU Legislative Meter System establishes a separate body called WELMEC (Western European Cooperation in Legal Metrology) for cooperation and agreement on the interdepartmental metrology field and has implemented the related system for each country in Europe for the mutual certification in the MID system. The WELMEC 7.2 [8] has been published for the evaluation and guidance of SW to meet the main, and it has been introduced into the MID certification system within the EU. Evaluation is conducted according to the rating required for the evaluation of the product through the risk management analysis.

2.2 OIML

The International Organization for Standardization metrology (Organisation Internationale de Metrolgie Legale) has internationally solved the problems of the administrative technology caused by the use of meters and developed international recommendations and international documents on legal metrology. Because it is based on information technology, the SW required for the legal meter is very important factor in the weighing and performance of the meter. In accordance with this trend, the

general requirement of the meter controlled by SW (OIML D-31 [9]) in the International Legal Metering Organization (OIML) was established as a reference standard.

When the meter is ready to download and update a new version of the SW remotely via communication, it need to check integrity. In the process legitimacy to the new SW is to go through the authorization process in the type approval authority. In addition, the meter records all events for the SW upgrade procedure by leaving an audit trail when the SW is upgraded.

3 Information Assurance Model

3.1 Objective

There is a need to set out the necessity of the following SW requirements through analysis of MID and OIML specification in Sect. 2.

Software Misuse/Fraud Prevention. The Meter's software must be capable of handling only the intended inputs. This requirement prevents the software from operating in unintended fashion or performing functions.

Accuracy of Stored Measurement Data. The instrument should be able to automatically store the measured values and confirm that the instrument has sufficient memory. It is ensured that the accuracy of the stored data is guaranteed by a mathematical method in order to prevent the data from being altered after being stored. In addition to instrumentation data, the instrument must record and protect events for access.

Stability of Instrumentation Software. The instrument should be able to operate the measurement function and not lose measurement data regardless of the network connection or its own hardware load.

Maintenance of Measurement Software. If you need to improve the functionality of the instrument's software, you can proceed with the software update. Software updates can be performed remotely, and the instrument itself can verify the update software.

Prevent Unauthorized Software. The instrument should be prevented from operating with unauthorized software. The instrument shall block illegal activities such as meter manipulation, illegal SW, parameter manipulation, and record relevant audit evidence.

3.2 Threat and Scenario

There is a need to avoid the software to act as an unintended procedure. It is a way to perform an unexpected function. Most of these issues require thorough exception handling by the developer. In this case, however, it cannot be filtered out in the development process because the developer would have to intentionally cheat (Tables 1 and 2).

Table 1. Software malfunction through input parameters

Requirement	Software misuse/fraud prevention
Vulnerability	Flow to unintended behavior not designed in software
Threat	Unexpected process progress with unintended input
Condition	**(At the development stage)** the occurrence of a software flow that was not designed due to developer's lack of exception handling **(At the development stage)** Developing intentional input parameters for special instructions or abnormal operation Then, if someone has access to the input parameters or input file when installing the power meter
Description	The instrument is cheated at the development stage before it is tested for type approval. Developers have special parameters to operate the instrument in illegal operating mode. - For yourself - Through someone's referral. This means that the problematic software is subjected to type approval testing.
Countermeasures	Branch check of procedure

Table 2. Infringement on the integrity and accuracy of stored data

Requirement	Accuracy of stored measurement data
Vulnerability	Accuracy threats to stored measurement data
Threat	**(SW)** Modify the storage of the software to tamper the data **(Terminal)** Access to the data file through the command input from the terminal
Condition	**(Data storage)** Storing data in plain text and lacking a method for verifying the data **(Malfunction Software)** Hacked software installed to perform data tempering when storing instrumentation values
Description	The data generated by the software's measurement functions must be reliable and accurate. The device must protect the stored data. The threats are as follows: - Tempering for stored data - Tempering at the time of storing data
Countermeasures	Mathematical method for ensuring the integrity of stored data

There is a need to prevent the tempering of the data stored in the measuring instrument. We must provide an environment in which users of the instrument can trust and use the data. To do so, there is need to define how to protect stored data from the software management requirements of the instrument.

There may be a vulnerability in the data storage phase, such as race conditions. A denial of service attack in a network environment can also be used to consume device resources. The measuring instrument must perform the measurement and data storage functions correctly under any conditions.

4 Testing Method and Future Work

Attempt to Unexpected Process Progress with Unintended Input. For this requirement we need to manage unexpected exceptions by input values. Power meter manufacturers may not develop their own software, or they may not release the source code. This is the easiest way to check branches through the source code, but this method is hardly realistic. Therefore, there is need to find the branch through the reverse engineering technique of the software installed in the measuring instrument. We need to build a testing environment that can detect branches and detect anomalies through reverse engineering of software.

Attempt to Cheat by Installing Unauthorized Software. There may be cases where the manufacturer operates the correct software in the device type approval and the cheat software is loaded in the operating state in the field. The device must record and thoroughly manage the changes made to the software using the software update log. (Software change subject/time/software unique value, etc.) This log is also as important as the measurement data. Detecting cheating through the event log is already done after the attack.

Attempt to Temper Measurement Data. The instrument should be able to automatically store the measured value. In this process, the device needs to make sure that it has enough space for data storage, and in the absence of free memory, it needs a policy to delete old data. Event logs and data files should be thoroughly protected and protected through the functionality of the operating system file system (Access control). The instrument should be able to record and block cases where external factors attempt to tamper with the data stored in the instrument, and there is need to identify how to check the integrity of the stored data through mathematical methods like hash function.

5 Conclusion

In order to ensure that these software control meters operate correctly, there is need to verify that the software is not tampered with and behaves correctly. Mechanical meters keep metering parameters such as gravitational acceleration, graduation intervals, and (the smallest unit covered by ‘scale’),

A software-controlled meter can change the weighing parameters after passing the type approval test, thereby making it possible to raise unfair profits. Therefore, there is a need for a method to confirm whether the program is the same at the time of the type approval test and in the field operation state. And before that, the software certification exam must be done thoroughly. In this study, we set five goals to be verified in the software certification test. And we wanted to design the threats that violated them and

how to achieve them. We are studying a technique for logically verifying that the software operation is correct, and there is need to develop a methodology for automatic testing through future studies.

Acknowledgments. 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. 20151210200080).

References

1. Van Gerwen, R., Jaarsma, S., Wilhite, R.: Smart metering. Leonardo-energy (2006)
2. Brophy Haney, A., Jamasb, T., Pollitt, M.G.: Smart metering and electricity demand: technology, economics and international experience (2009)
3. Ahmad, S.: Smart metering and home automation solutions for the next decade. In: 2011 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC), pp. 200–204. IEEE (2011)
4. Kabalci, Y.: A survey on smart metering and smart grid communication. Renew. Sustain. Energy Rev. **57**, 302–318 (2016)
5. Depuru, S.S.S.R., Wang, L., Devabhaktuni, V., Gudi, N.: Smart meters for power grid—challenges, issues, advantages and status. In: 2011 IEEE/PES Power Systems Conference and Ex-position (PSCE), pp. 1–7. IEEE (2011)
6. Thiel, F., Hartmann, V., Grottker, U., Richter, D.: IT security standards and legal metrology–transfer and validation. In: EPJ Web of Conferences, vol. 77, p. 00001. EDP Sciences (2014)
7. Bernieri, A., Betta, G., Ferrigno, L., Laracca, M., Moriello, R.S.L.: Electrical energy metering: some challenges of the European directive on measuring instruments (MID). Measurement **46**(9), 3347–3354 (2013)
8. WELMEC Guide 7.2: Software Guide (Measuring Instruments Directive 2004/22/EC)
9. Organisation Internationale de Métrologie Légale (OIML), General requirements for software controlled measuring instruments, OIML D-31 (2008)

Extraction of Camera Parameters for Image-Based Motion Capture

Nu-lee Song, Man-ki Kim, Hern-soo Han,
Dong-ho Kim, and Gye-young Kim(✉)

378, Sangdo-ro, Dongjak-gu, Seoul, Republic of Korea

Abstract. There are some limitations in acquiring motion data from images without markers. But, low-cost ultra-speed cameras have been commercialized, therefore we can overcome the limitations using ultra-speed cameras. In order to obtain motion data from images, three or more cameras are needed and the established information, that is, camera parameters for all cameras are known. It is called camera calibration. This paper presents an efficient method for camera calibration based on a calibration pattern which consists of a circle and a rectangle.

Keywords: Camera calibration · Motion capture · External parameter

1 Introduction

Camera Calibration, which extracts viewing parameters of a camera, is an essential procedure for developing an image based motion capture. Haralick proposed how to use the 2D perspective projection of a rectangle of unknown size and position in 3D space [1]. The method is very simple and efficient in the respect of computational complexity. But it assumes that the focal length of a camera should be known. What we know the focal length is also a calibration problem. Therefore, the method is incomplete. Han and Rhee proposed a method using a specially designed circular pattern with two internal dots, a center dot and a direction dot [2]. The method is also simple and efficient. But it assumes that the optical axis of a camera is directed to the origin of a world coordinate system. In accordance with the assumption, the camera's view is adjusted manually until the center dot appears at the image center. This work is very tedious and inaccurate. Zhang used a chessboard pattern to calculate intrinsic parameters of a camera [3]. Miyagawa proposed from five points, whose 3D world coordinates are known, on two orthogonal 1-D objects [4]. Chen utilized five-point correspondences between image and world coordinates [5]. In this paper, we propose a new method using properties of a circle and a rectangle contained in a calibration pattern for image-based motion capture.

2 Calibration Pattern

The calibration pattern developed in this letter is consisted of a rectangle, a circle, and a center dot as shown in Fig. 1.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 184–188, 2020.
https://doi.org/10.1007/978-981-13-9341-9_31

Let (X, Y, Z) and (x, z) are the eye coordinate system and the image coordinate system, respectively. The rectangle and the circle have known width W, length L, and diameter R. We assume that the camera lens is the origin and that the lens views down the y axis. The image plane is an unknown distance in front of the lens and is orthogonal to the optical lens axis. A calibration pattern lies in the plane. In this model, the pan angle of a camera represents the angle of counter clockwise rotation around Z axis. The tilt angle and swing angle represent the rotations around Y axis and X axis, respectively. If the two angles, tilt and pan, are not zero, the circle and the rectangle in 3D space, the eye coordinate system, will be appeared an ellipse and a trapezoid in 2D image, the image coordinate system.

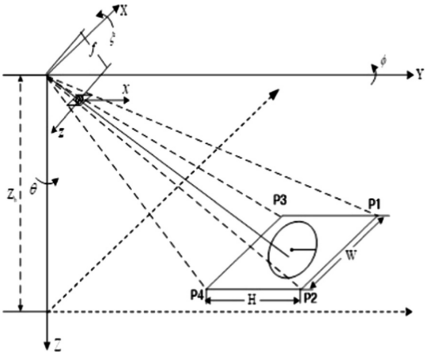


Fig. 1. Calibration pattern, world and image coordinate system (Left)

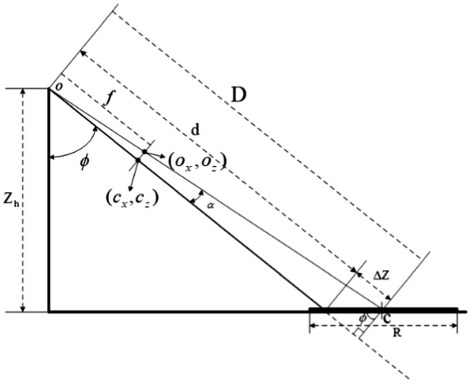


Fig. 2. The geometric relations of the circle and the corresponding image (Right)

We can assume that the corners of the rectangle are given by

$$P_1 = \begin{pmatrix} X_1 \\ Y_1 \\ Z_1 \end{pmatrix}, P_2 = \begin{pmatrix} X_1 + W \\ Y_1 \\ Z_1 \end{pmatrix}, P_3 = \begin{pmatrix} X_1 \\ Y_1 + L \\ Z_1 \end{pmatrix}, P_4 = \begin{pmatrix} X_1 + W \\ Y_1 + L \\ Z_1 \end{pmatrix} \quad (1)$$

Where $Y_1 > f$, and that the corresponding perspective projection of these corners are

$$P_1^* = \begin{pmatrix} X_1^* \\ Z_1^* \end{pmatrix}, P_2^* = \begin{pmatrix} X_2^* \\ Z_2^* \end{pmatrix}, P_3^* = \begin{pmatrix} X_3^* \\ Z_3^* \end{pmatrix}, P_4^* = \begin{pmatrix} X_4^* \\ Z_4^* \end{pmatrix} \quad (2)$$

We assume that the lengths of the short axis and the long axis of the ellipse are r' and r'', respectively.

3 Method for Camera Calibration

By referring Haralick [1], we can first drives the following Eqs. (3) to (5), using four corner points of the rectangle and the principle of a perspective transformation to solve for φ, θ, ξ with λ_1, λ_2, λ_3, λ_4, W and L all unknown.

$$\phi = \tan^{-1}\frac{\left(x_2'z_4' - x_4'z_2'\right)\left(z_1' - z_3'\right) - \left(x_1'z_3' - x_3'z_1'\right)\left(z_2' - z_4'\right)}{f\left[\left(x_1' - x_3'\right)\left(z_2' - z_4'\right) - \left(x_2' - x_4'\right)\left(z_1' - z_3'\right)\right]} \tag{3}$$

$$\theta = \tan^{-1}\frac{cos\phi\left[\left(x_1'z_3' - x_3'z_1'\right)\left(z_2' - z_4'\right) - \left(x_2'z_4' - x_4'z_2'\right)\left(z_1' - z_3'\right)\right]}{f\left[\left(x_1' - x_3'\right)\left(z_2' - z_4'\right) - \left(x_2' - x_4'\right)\left(z_1' - z_3'\right)\right]} \tag{4}$$

$$\xi = \tan^{-1}\frac{\left[A\left(z_1^* - z_3^*\right) - B\left(z_2^* - z_4^*\right) - C\left(z_1^* - z_2^*\right) - D\left(z_3^* - z_4^*\right)\right]}{\left[A\left(x_1^* - x_3^*\right) - B\left(x_2^* - x_4^*\right) - C\left(x_1^* - x_2^*\right) - D\left(x_3^* - x_4^*\right)\right]} \tag{5}$$

Where

$$\begin{pmatrix} x_i' \\ z_i \end{pmatrix} = \begin{pmatrix} cos\xi & -sin\xi \\ sin\xi & cos\xi \end{pmatrix}\begin{pmatrix} x_i^* \\ z_i^* \end{pmatrix}, \quad i = 1,2,3,4$$

$$\mathrm{A} = \frac{x_2^*z_4^* - x_4^*z_2^*}{E},\ \mathrm{B} = \frac{x_1^*z_3^* - x_3^*z_1^*}{E},\ \mathrm{C} = \frac{x_3^*z_4^* - x_4^*z_3^*}{F},\ \mathrm{D} = \frac{x_1^*z_2^* - x_2^*z_1^*}{F}$$

$$\mathrm{E} = f\left[\left(x_1^* - x_3^*\right)\left(z_2^* - z_4^*\right) - \left(x_2^* - x_4^*\right)\left(z_1^* - z_3^*\right)\right]$$
$$\mathrm{F} = f\left[\left(x_1^* - x_2^*\right)\left(z_3^* - z_4^*\right) - \left(x_3^* - x_4^*\right)\left(z_1^* - z_2^*\right)\right]$$

Then, we can derives the following Eqs. (6) to (8), to determine Z_1, X_1 and Y_1 under the assumption that the sides, W and L, of the rectangle are given.

$$\mathrm{Z}_1 = \frac{\left(fsin\phi + z_3'cos\phi\right)\left(fsin\phi + z_1'cos\phi\right)L}{f\left(x_3' - x_1'\right)sin\theta sin\phi + f\left(z_1' - z_3'\right)cos\theta + \left(x_3'z_1' - x_1'z_3'\right)sin\theta sin\phi} \tag{6}$$

$$\mathrm{X}_1 = Z_1\frac{x_1'cos\theta - fsin\theta cos\phi + z_1'sin\theta sin\phi}{fsin\phi + Z_1'cos\phi} \tag{7}$$

$$\mathrm{Y}_1 = Z_1\frac{x_1'sin\theta + fcos\theta cos\phi - z_1'cos\theta sin\phi}{fsin\phi + Z_1'cos\phi} \tag{8}$$

At this point, we note that the above Eqs. (3) to (8), are derived from the known focal length f and the known side length L. That is, if we don't know f and L, we can't uniquely determine the viewing parameters of the camera. It is easy to know L, but it is very difficult to know f. We propose a new camera calibration method with an unknown focal length.

We can rewrite (7) as Eq. (9). In (9), L_{f_i} and Z_{f_i} denote an estimated length of L and an estimated length of Z_1 with a given focal length f_i, respectively.

$$\mathrm{L_{f_i}} = Z_{f_i}\frac{f_i\left(x_3^{'}-x_1^{'}\right)sin\theta sin\phi + f_i\left(z_1^{'}+z_3^{'}\right)cos\theta + \left(x_3^{'}z_1^{'}-x_1^{'}z_3^{'}\right)sin\theta cos\phi}{\left(f_i sin\phi + z_3^{'}cos\phi\right)\left(f_i sin\phi + z_i^{'}cos\phi\right)L} \tag{9}$$

We explain how to compute the Z_{f_i} with the circle in the calibration pattern. We can drive Eq. (10) form the geometric relations between the circle and the corresponding image shown in Fig. 2. In (10), $r = r'$ if the angle between the short axis of the ellipse and the z axis is greater than the angle between the long axis of the ellipse and the z axis, otherwise $r = r''$. (c_x, c_y) are the center coordinates of the image plane and (o_x, o_y) denote the image coordinates of the center dot.

$$\mathrm{Z_h} = d \cdot cos\phi \tag{10}$$

where

$$\mathrm{d} = \mathrm{D} - \Delta\mathrm{x} - \Delta\mathrm{z},\ \Delta\mathrm{z} = \mathrm{D}\cdot\tan\alpha\cdot\tan\phi,\ \Delta\mathrm{x} = \mathrm{D}\cdot\tan\beta\cdot\tan\theta$$

$$\mathrm{D} = \frac{\mathrm{f_i}\cdot R\cdot cos\phi}{r},\ \alpha = \tan^{-1}\left(\frac{O_z - C_z}{f_i}\right)\ and\ \beta = \tan^{-1}\left(\frac{O_x - C_x}{f_i}\right)$$

Now, we can determine the estimated values, L_{f_i} and Z_{f_i}, using (9) and (10) for a given f_i. Unfortunately we don't have the real focal length f of the camera. But we can select a focal length f_i minimizing the difference between L and L_{f_i} as Eq. (11). In (11), $f_{\min}$ and $f_{\max}$ denote the minimum and the maximum focal length of the camera. The number of f_i s is N. So, the distance Δf between f_i and f_{i+1} is $(f_{\max} - f_{\min})/N$. N is experimentally determined.

$$\begin{aligned}&\mathrm{Min_{diff}} = \min\left\{\mathrm{Diff}_{f_i} | f_i = f_{min}..f_{max}\right\}\\ &\mathrm{f_c} = argmin\left\{\mathrm{Diff}_{f_i} | f_i = f_{min}..f_{max}\right\}\\ &\text{where } \mathrm{Diff}_{f_i} = \left|L - L_{f_i}\right|\end{aligned} \tag{11}$$

For the purpose of increasing the accuracy of the viewing parameters from camera calibration, we iteratively apply Eq. (11) to the new range from $f_{\min} = f_c - \Delta f/2$ to $f_{\max} = f_c + \Delta f/2$ during Min_{diff} is improved.

4 Experimental Results

In order to implement the proposed method and evaluate the performance, Intel Core I7-7700 (3.6 GHz) and input images are acquired by CANON CMOS camera mounted on PAN-TILT unit which can changes the FOV of a camera. The acquired images are 768 × 494 in size with 32 bits per pixel. The accuracy of the viewing parameters from camera calibration can be evaluated by comparing the calculated position of an object

with its true positional value. An object is placed at known coordinate position in the world coordinate system and its corresponding positions are calculated using the calibrated camera.

The real size of the rectangle of the pattern used for the performance evaluation is 100 × 100 (mm) and the diameter of the circle is 80 (mm). The distance between two neighbor points is 50 (mm). We calculate the focal lengths and the calibration errors with varying each viewing parameter in order to evaluate the accuracy and the robustness. Table 1 show the calibration error. The results are obtained from 150 experiments.

Table 1. Calibration errors

	Calibration error (%)			
	Minimum	Maximum	Average	Standard deviation
Height	1.86	3.40	3.09	0.18
Swing	1.21	3.29	2.31	0.25
Pan	2.04	3.68	2.90	0.21
Tilt	0.90	3.25	1.29	0.44
Integrated	1.86	3.44	2.72	0.28

5 Conclusion

We have presented a new calibration method using the properties of a circle and a rectangle and shown the performance evaluation results in respect with stability and calibration error. According to the results, we can know that the standard deviation of the determined focal lengths and the average of calibration error are about 2(%) and 2.72(%).

Acknowledgements. This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2018-2018-0-01419) supervised by the IITP (Institute for Information & communications Technology Promotion).

References

1. Haralick, R.M.: Determining camera parameters from the perspective projection of a rectangle. Pattern Recogn. **22**(3), 225–230 (1989)
2. Han, M.-H., Rhee, S.: Camera calibration for three-dimensional measurement. Pattern Recogn. **25**(2), 155–164 (1992)
3. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Mach. Intell. **22**(11), 1330–1334 (2000)
4. Miyagawa, I.: Simple camera calibration form a single image using five points on two orthogonal 1-D objects. IEEE Trans. Image Process. **19**(6), 1528–1538 (2010)
5. Chen, H.-T.: Geometry-based camera calibration using five-point correspondences from a single image. IEEE Trans. Circ. Syst. Video Technol. **27**(12), 2555–2566 (2017)

Serverless Framework for Efficient Resource Management in Docker Environment

Sangwook Han[1], Minsu Chae[1], and HwaMin Lee[2(✉)]

[1] Department of Computer Science,
Soonchunhyang University, Asan, South Korea
{sanguk, cms}@sch.ac.kr
[2] Department of Computer Software Engineering,
Soonchunhyang University, Asan, South Korea
leehm@sch.ac.kr

Abstract. APIs provided by the Docker is executed through the container engine. Thus, it is a reality that the speed of creation and deletion of containers or requests for information is very slow. The load is even worse when many containers sending API requests at once. In this paper, we propose a method to apply module related to resource management in a container to reduce the load on API request generated in serverless environment. And we propose a new framework that manages resources of several containers or multiple Docker serves.

Keywords: Docker · Serverless framework · Resource management

1 Introduction

Recently, data centers are virtualization the cloud, not just servers, but also storage and network infrastructure. Therefore, hypervisor-based virtual machine technologies such as Xen, KVM, etc. are mainly used to build cloud infrastructure. However, hypervisor-based virtual machine creation technologies such as Xen or KVM require very high resource usage when creating virtual machines, and waste resources when not using allocated resources [1].

Serverless method using Docker is very useful when temporary service like web server is needed, and it is advantageous for program developer to focus on logic without worrying about development environment such as API. Therefore, it is expected that future cloud environment will be more convenient for users by using a container method that provides the convenience of resource management like Docker.

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-2017R1A2B4010570) and the Ministry of Science and ICT, Korea, under the ITRC (Information Technology Research Center) support program (IITP-2017-2015-0-00403) supervised by the IITP (Institute for Information & communications Technology Promotion).

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 189–194, 2020.
https://doi.org/10.1007/978-981-13-9341-9_32

When a large number of users request a service, the efficiency of the container method is high when the method of using the container is compared with the method of using the existing hypervisor [2]. Therefore, a framework that can support efficient scheduling by helping cloud users complete a work quickly is very important [3].

2 Related Research

2.1 Docker

Docker is a container-based application automation open source project. Docker uses the virtualization features provided by the Linux kernel and generates containers by changing over from execution files, such as code, runtime, and system libraries to the image. Unlike virtual machines, Docker does not require a separate operating system. The Docker shares the kernel of the container but runs in isolation from the user space on the host operating system. Previously, it was only available on Linux, but since Windows 10, you can use the Docker with Hyper-V functionality [4].

2.2 Serverless

Serverless is an environment that allows application providers to use without considering the server. It works on the server only when the application is running. Currently, there are many platforms based on Serverless [5], and Serverless [6, 7] is being studied in various environments. Previously, it was based on event-driven processing, but it is being used as a container management and software development strategy in recent years. In the case of Serverless Computing implemented by McGrath [8], the maximum memory size is not limited to the required memory size, and the Docker API is used. When using the API for resource monitoring, API requests are made to the Docker engine to obtain one container information. At this time, since the Docker engine fetches the information of the active container, performance deterioration occurs. Therefore, there is a problem that the resource efficiency is lowered.

Table 1. The data structure used in this paper

Variable	Description
CONTAINERS	Number of containers to monitor for resources
$TIME_{resource}$	Time spent monitoring resources
$TIME_{network}$	Time spent on network transmission

Table 1 defines the time required for Docker resource monitoring. The time required for resource monitoring based on mathematical model [9] is as follows.

$$TIME_{resource} + TIME_{network} \quad (1)$$

However, the Docker API can only know one container's resources in a single request. The following formula is the time required by the number of requests in the container.

$$\text{CONTAINERS} * \text{TIME}_{\text{resource}} + \text{CONTAINERS} * \text{TIME}_{\text{network}} \quad (2)$$

As the number of containers to monitor the resource grows, the problem becomes longer.

In order to provide a convenient user interface environment for the general user, an agent system that can perform the user's work beyond the current Windows-based user interface should be provided. In addition, research and development of a middleware platform that can be applied for extension and use of agent system service should be done.

3 Serverless Framework Design

3.1 Docker-Based Serverless Framework

Currently, the Docker-based serverless environment is configured as shown in Fig. 1. As shown in Fig. 1, our model consists of containers that provide each service and servers that support applications or the web. Management module container requests data of creation, deletion, resource information, etc. of service container to Docker engine to perform specific tasks. At this time, a load is generated. Therefore, this paper proposes a framework to solve this problem.

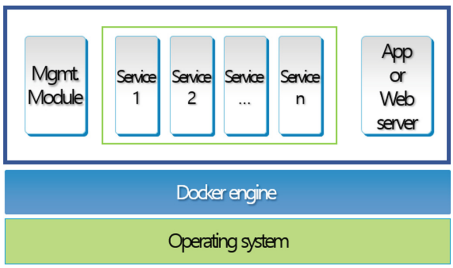


Fig. 1. Docker-based serverless architecture

Figure 2 shows the proposed Docker-based serverless structure. A module for monitoring resources is embedded in each service container, and the host resource management monitoring module receives the container information directly from the service container without requesting API to the container engine. Changing the structure in this way is also very efficient for managing multiple Docker servers.

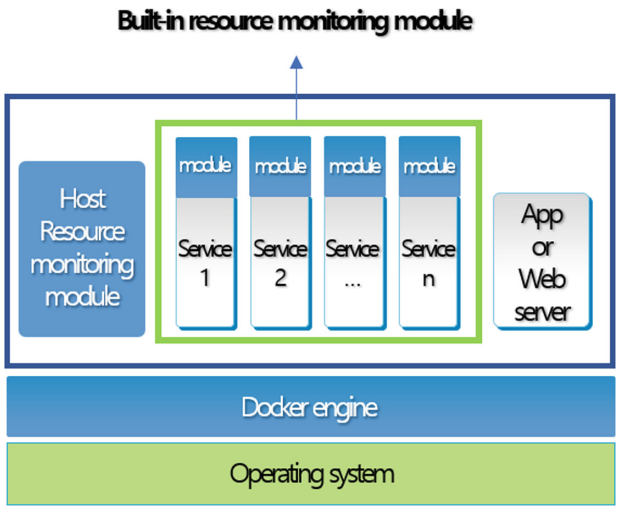


Fig. 2. Proposed Docker-based serverless architecture.

By using the proposed architecture, it is possible to efficiently manage resource distribution scheduling according to user ‘s demand and a plurality of Docker servers. Figure 3 shows our proposed serverless framework. In order to manage multiple Docker servers with the proposed structure as shown in Fig. 3, it is possible to design a whole host module that can eliminate the existing performance degradation and can perform quick resource inference and distribution. A detailed description of each module is given in Sect. 3.2.

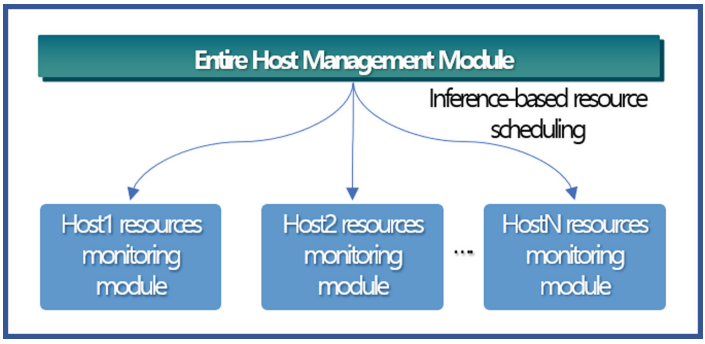


Fig. 3. Our serverless framework

3.2 Docker Resource Monitoring Module

Since there is a problem of performance degradation when processing using the Docker API, a module for measuring CPU utilization and memory usage should be

developed to solve the performance degradation problem. First, resource monitoring modules for containers and servers are essential. In order to reduce the performance degradation due to the network, the result collected in the container is transmitted to the monitoring module of the server that executes the container. The monitoring module of the server transmits the resource monitoring information and the resource monitoring information of the containers performed in the server to the management module.

3.3 Resource Inference Module

In order to efficiently manage resources, we deduce resources based on the type of service being executed and the image that has been executed. To do this, we analyze the resource usage according to each service type and use the information collected by the container resource monitoring module to analyze the performance based on the previously executed image.

The source code below represents the resource reasoning module algorithm.

```
function getResourceInference(type, image){
  cpu = getEstimateCPUusageByServiceType(type)
  memory = getEstimateMemoryuageByServiceType(type)
  cpu2 = getEstimateCPUusageByMonitoring(image)
  memory2 = getEstimateMemoryusageByMonitoring(image)
  needCPU = (cpu + cpu2) / 2
  needMemory = (memory + memory2) / 2
  return needCPU, needMemory
  }
```

3.4 Entire Host Management Module

The entire host management module collects the resource information of the Docker servers using the resource monitoring module, and allocates the task requested by the user based on the collected information. Manage resource efficiently using resource inference module.

The source code represents the resource allocation algorithm of the entire management module.

```
function containerAllocation(hosts, type, image){
  cpu, memory = getResourceInference(type, image)
  select = getUselessHost(hosts)
  for(host : hosts)
    if(host.isAllocation(cpu, memory))
      if(select.addUsage(cpu, memory)>host.addUsage(cpu,
memory)
        select = host
  container = select.createContainer(image)
   return container
 }
```

4 Conclusion

In this paper, we design a serverless framework for efficient resource management in Docker environment. The designed framework has the following effects. First, by using container, developers can focus on logic and develop. Second, resource monitoring does not cause performance degradation because it is not an existing Docker API communication. Third, efficient scheduling and resource management is possible through resource inference. We plan to implement the designed framework in the future. At present, only the resource allocation algorithm considering only resource utilization is implemented, but it is planned to design an improved resource allocation algorithm in the future.

References

1. Bae, Y.-M., Jung, S.-J., Soh, W.-Y.: Comparative analysis of the virtual machine and containers methods through the web server configuration. J. Korea Institute Inf. Commun. Eng. **18**(11), 2670–2677 (2014)
2. Chae, M., Lee, H., Lee, K.: A performance comparison of linux containers and virtual machines using Docker and KVM. Cluster Comput., 1–11 (2017)
3. Moon, Y., et al.: A slave ants based ant colony optimization algorithm for task scheduling in cloud computing environments. Hum.-Centric Comput. Inf. Sci. **7**(1), 28 (2017)
4. Merkel, D.: Docker: lightweight linux containers for consistent development and deployment. Linux J. **2014**(239), 2 (2014)
5. Lynn, T., et al.: A preliminary review of enterprise serverless cloud computing (Function-as-a-Service) platforms. In: 2017 IEEE International Conference on Cloud Computing Technology and Science (CloudCom). IEEE (2017)
6. de Lara, E., et al.: Hierarchical serverless computing for the mobile edge. In: IEEE/ACM Symposium on Edge Computing (SEC). IEEE (2016)
7. Dash, S., Dash, D.K.: Serverless cloud computing framework for smart grid architecture. In: 2016 IEEE 7th Power India International Conference (PIICON). IEEE (2016)
8. McGrath, G., Brenner, P.R.: Serverless computing: design, implementation, and performance. In: 2017 IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW). IEEE, (2017)
9. Respondek, J.: Numerical approach to the non-linear diofantic equations with applications to the controllability of infinite dimensional dynamical systems. Int. J. Control **78**(13), 1017–1030 (2005)

IoT Based Management System for Livestock Farming

Meonghun Lee[1], Haengkon Kim[2], Ha Jin Hwang[3], and Hyun Yoe[4(✉)]

[1] Department of Agricultural Engineering, NAAS, Jeonbuk 55365, Republic of Korea
leemh5544@gmail.com
[2] School of Information Technology Engineering, Catholic University of Daegu, Gyeongbuk 38430, Republic of Korea
hangkon@cu.ac.kr
[3] Sunway University Business School, Sunway University, Subang Jaya, Malaysia
hjhwang@sunway.edu.my
[4] Department of Information and Communication Engineering, Sunchon National University, Jeonnam 57922, Republic of Korea
yhyun@sunchon.ac.kr

Abstract. Management system for livestock farming is very complex system in general and thus is strictly influenced by several constrain such as climate, weather, and livestock conditions, and so on. Livestock managers (Farmers) try to optimize utilization of agricultural resources to obtain good product having both quality and quantity. Particularly, various environmental IoT sensors are being used to substitute the areas that are difficult for people to do directly. Although the Korea livestock industry is becoming larger in scale and systematic, it is still receiving damages from various diseases and difficulty in maintaining growth and development environment. To reduce such damages, this paper proposes an IoT based management system for collecting environmental information for real-time monitoring by using various environmental IoT sensors. The proposed system allows manager to easily monitor via PC and smart devices by deviating from the existing manual method of directly checking growth and development environment. In addition, it has been enhanced to extend the life of IoT sensors through the method of efficiently collecting energy data for extending the slip time of IoT node while minimizing routing path.

Keywords: IoT · Livestock · Management system · Routing

1 Introduction

To reduce natural damages and the incidence of diseases fatal to livestock and stably maintain productivity, it is necessary to create optimal growth environment for which monitoring system for checking livestock barn environment is needed. To solve such issues, this paper proposes a IoT based management system for livestock farming [1, 2].

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 195–201, 2020.
https://doi.org/10.1007/978-981-13-9341-9_33

The proposed system consists of various environmental IoT sensors and mobile technology, and it is a system that collects the changing environmental information of livestock barn via environmental IoT sensor and transmits it wirelessly for analysis to notify the current livestock barn situation to managers. It is expected this system will allow livestock barn environment monitoring in real-time, active respond to different situations and prevention of damages of livestock farms in emergency situations.

2 System Design

2.1 System Architecture

In this paper, an IoT based Management System for Livestock Farming was designed through environmental IoT sensor information based on wireless sensor network. The proposed system deduced components needed to maintain livestock barn environmental condition through which the architecture of IoT based environmental monitoring system was designed. The system components mainly consist of physical layer, middle layer and application layer. Figure 1 is system architecture based on system components. The physical layer consists of various environmental IoT sensors needed for livestock barn environment monitoring and the middle layer consists of environmental information manager and management server, web server, app server and database. The application layer consists of livestock barn environment monitoring service, abnormal situation notification service, PC web, smart phone app and infrastructure to exchange information with manager.

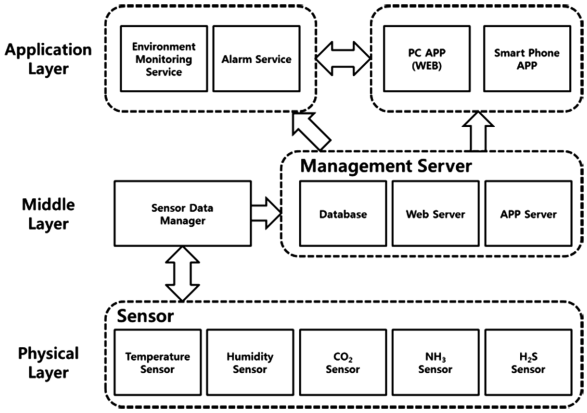


Fig. 1. IoT based management system architecture

2.2 System Configuration

IoT based Management System can measure livestock barn environment and collect data in real-time by implement wireless sensor network technology and environmental IoT sensor in livestock barn. This system was developed by dividing the sensors to be

used for livestock barn environment monitoring into five types. Sensors for temperature, humidity, CO_2, NH_3 and H_2S closely related to livestock barn growth environment were used to design a more precise environment monitoring system. In addition, it consists of database for storing and analyzing environmental information, management server and PC and smart device for monitoring by manager.

2.3 System Process

The proposed system operates through the operational process as shown in Fig. 2. First, temperature, humidity, CO_2, NH_3 and H_2S data are collected at certain intervals through the environmental IoT sensors in livestock barn to transmit it to environmental information manager. The environmental information manager processes the received information and stores it in the database, and the management server runs analysis based on this. Accordingly, it provides environmental monitoring information and notification service during abnormal situation. In addition, manager can monitor at remote place through PC or smart device.

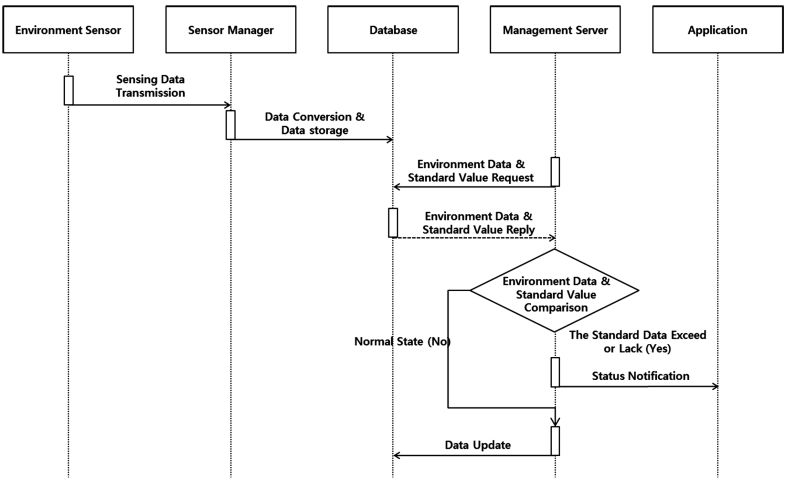


Fig. 2. IoT based management system process

2.4 Energy Efficient Data Collection Method

When collecting data through connected sensor via wireless sensor network, existing method entailed significant energy consumptions with no efficiency [3]. Accordingly, this paper proposes an energy-efficient way of collecting data needed for livestock barn environment monitoring. In the proposed method, cluster header is used in wireless sensor network environment and cluster header node combines the data of cluster member node to transmit to sink, thereby reducing the amount of communication between nodes. Since energy lifespan of IoT node shortens when amount of communication increases in addition to frequent routing path setting, slip time of IoT node is

extended as cluster head adjusts the schedule of member node based on TDMA schedule. In the existing LEACH study, there was a limitation of having to know the energy of every IoT node but this method only uses the information of corresponding IoT node in the selection of cluster head [4–6]. Accordingly, it is not necessary to know the energy of every node when selecting cluster head. Network resources can be efficiently used through the proposed method, while also minimizing the energy consumption of IoT node.

3 Implementation

3.1 Test-Bed

The IoT based Management System proposed in this paper consists of IoT node to collect livestock barn environment information and wireless sensor network technology was applied to send in real-time environmental data to remote location to allow manager to conveniently monitor livestock barn environment. To verify the system, the system was placed on test bed as shown in Fig. 3.

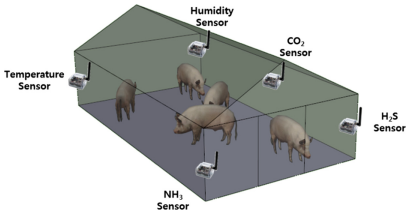


Fig. 3. IoT based Livestock Management System

3.2 Environmental Analysis

In the test-bed, each environmental IoT sensor was placed according to area. Considering the size of the test bed, sensors were in livestock barn by dividing the area into four areas for accurate measurement. As for the environmental data collection period, it was set as 48 h and the analysis of collected data is as follows. First off, data from temperature sensor are as shown in Fig. 4. In terms of temperature, there is a need to lower temperature during the day through ventilation or air conditioning unit based on optimal growth condition, and raise temperature at night through heating as temperature drastically falls. Humidity was consistent in average but there is a need to control humidity when needed by using ventilation and water. Appropriate humidity for breeding pig is 60–65% and humidity should be controlled since respiratory disease can occur when humidity falls. In terms of CO_2, it showed the result of increasing and decreasing according to temperature. In the case of heating to raise temperature, CO_2 amount increases drastically. Although CO_2 is not harmful gas with direct threat to livestock, livestock can die as amount of oxygen decreases. In addition, CO_2 becomes a

good index for examining the quality of air in livestock barn. NH_3 is main gas that causes the health issue of livestock.

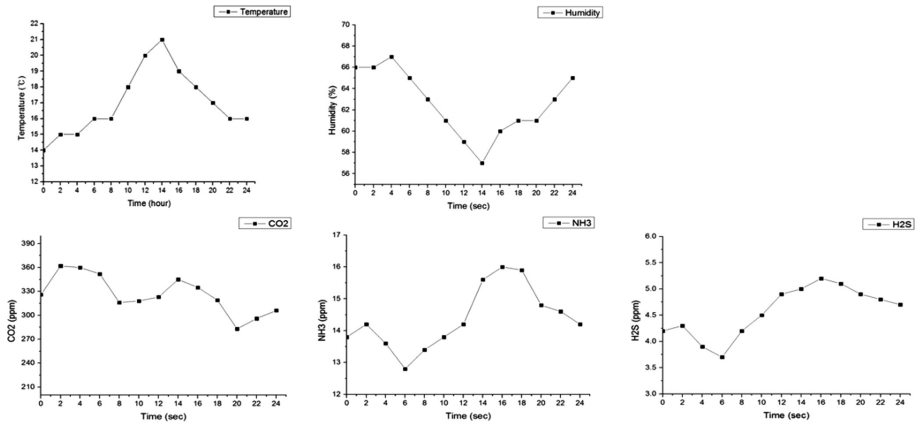


Fig. 4. Measured graph (Temperature, Humidity, CO_2, NH_3 and H_2S)

High concentration level ammonia causes the symptoms of headache, vomit and loss of appetite and in the case of inhaling it for a long period of time it causes damage to cilium that removes foreign object in airway. H_2S can occur from livestock excretion. Hydrogen sulfide is colorless toxic gas with odor. Since it is harmful gas that could cause momentary unconsciousness or death even with few respirations when its concentration surpasses 700 ppm, it must be controlled through monitoring. However, environment does not change immediately but becomes applied gradually even when measures are taken when every environment condition increased and decreased. Accordingly, it should be able to predict this and take according measure for prevention. This requires the accumulation of environmental data and there is a need for standard on when to provide notification service through mean values.

3.3 Data Collection Method

The proposed method was to reduce the amount of communication between nodes by cluster header node that combines the data of cluster member node and transmit them to sink using cluster header. In addition, it was confirmed that network resources are efficiently used and energy consumption of IoT node is minimized by extending the slip time of node. Figure 5 is a graph of IoT nodes life results in comparison between existing LEACH, TAG method and the proposed method. The experiment result showed the IoT nodes life increased by about 26%.

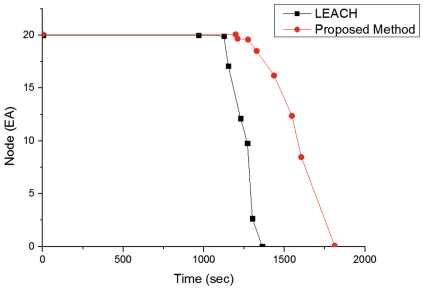


Fig. 5. Comparison of IoT nodes life

4 Conclusions

This paper proposed an IoT based Management System for Livestock Farming. To obtain accurate livestock barn environment information, various environmental IoT sensors and efficient data collection method were applied to implement a system more efficient than existing method. In addition, the system performance was confirmed through the verification using test bed, and it is expected that more precise monitoring will be possible in the future through accumulated data. It is expected through the proposed system that eco-friendly agriculture and low-cost & high-efficiency scientific stockbreeding will be realized by monitoring various environment factors that have important effects in forming optimal livestock barn growth environment such as temperature, humidity and various gases.

Acknowledgments. This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2018-2013-1-00877) supervised by the IITP (Institute for Information & communications Technology Promotion).

References

1. Lee, M., Kim, H., Yoe, H.: Intelligent environment management system for controlled horticulture. In: 2017 4th NAFOSTED Conference on Information and Computer Science, Hanoi, pp. 116–119 (2017)
2. Jeong, H., Lee, M., Yoe, H.: An environment monitoring system for livestock barn based on energy efficient wireless sensor network. Asian Int. J. Life Sci. (Supplement 11), 679–692 (2015)
3. Zhihui, H.: Research on WSN routing algorithm based on energy efficiency. In: 2015 Sixth International Conference on Intelligent Systems Design and Engineering Applications (ISDEA), Guiyang, pp. 696–699 (2015)
4. Singh, K.: WSN LEACH based protocols: a structural analysis. In: 2015 International Conference and Workshop on Computing and Communication (IEMCON), Vancouver, BC, pp. 1–7 (2015)

5. Sumitha, D.H., Kumar, M.V., Priya, M., Mangai, M.A.: Energy optimizion in leach algorithms using NS2. In: 2014 International Conference on Electronics and Communication Systems (ICECS), Coimbatore, pp. 1–3 (2014)
6. Pawar, S., Kasliwal, P.: Design and evaluation of En-LEACH routing protocol for wireless sensor network. In: 2012 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, Sanya, pp. 489–492 (2012)

Automatic Segmentation of Human Spine with Deep Neural Network

Xu Yin, Yan Li, and Byeong-seok Shin(✉)

Department of Computer Engineering, INHA University, Incheon, South Korea
22181731@inha.edu, {leeyeon, bsshin}@inha.ac.kr

Abstract. Considering that a CT scan produces cross-sectional images of specific areas of a scanned object, medical images of the spine are much more complex than those of other organs. In this study, the entire task of CT segmentation is viewed as a binary classification problem. We modify two deep learning networks—U-Net and SegNet—that are often used in the field of semantic segmentation. We made following modifications. Firstly, considering the size of CT images, we further reduce the number of convolutional layers. Then we use an element-wise method instead of concatenation in U-Net. Lastly, we select a new loss function as an evaluation criterion. According to the experimental results, we conclude that U-Net is not applicable when the training set is large, in which case we cannot prevent overfitting. However, SegNet performs better than U-Net in CT images segmentation.

Keywords: Semantic segmentation · Neural network · Medical images

1 Introduction

Convolutional neural network [1, 2], the most popular deep neural network, is now mainly used in computer vision. Related sub-fields include image classification and semantic segmentation. In medical imaging, in particular, doctors have a clear understanding of diseases when objects are segmented or further classified from original images. It should be noted that medical images are generated using different imaging technologies, such as computed tomography (CT), and magnetic resonance imaging (MRI). Both of these are applied to depict the structure of the human body, including blood vessels, bones, and organs. Segmentation is mainly used for detection (of unhealthy tissue) and depiction.

Early researchers including Dan et al. [3] tried to segment electron microscopy (EM) images. They applied classifiers to each pixel with a sliding window method and then extracted a patch around a pixel. This pixel-based method performed well in some cases, but it had several drawbacks. It consumed too much time for training, and it seemed difficult to learn the high-level features. Similar to FCN [4], Ronneberger et al. [5] designed a U-shaped architecture (U-Net). They considered that the expansive path of features would be more or less symmetric to the contracting path, so they added more feature channels in the upsampling stage. Feature maps of the last layer are concatenated to the previous ones, and finally a convolutional filter size of 1×1

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 202–207, 2020.
https://doi.org/10.1007/978-981-13-9341-9_34

produces segmentations. At the same time, U-Net divides images into regions and uses data augmentation to increase the number of training sets.

In this paper, we use similar architectures to SegNet [6] and U-Net to complete a medical image segmentation job. The remaining sections of this paper are organized as follows: in Sect. 2, we design two networks, based on U-Net and SegNet. In Sect. 3, we present a comparison between the two networks. In Sect. 4, we summarize our work.

2 Deep Network Design Based on U-Net and SegNet

In order to segment the spine region from a cross-sectional spine CT image [3, 7], the U-Net and SegNet networks introduced in the previous section were applied for segmentation tasks. The total number of images in the data set is 916 (including 716 segmented manually in ImageJ, an open source image processing software designed for scientific multidimensional images). The original size of images is 512 × 512. The greatest difference between classification and segmentation is that we should do a prediction task for each pixel, resulting in an output size of 512 × 512.

Based on SegNet, the most important point in this architecture is that each pooling layer saves the corresponding indices, except for location information. Although it appears to be a hierarchy system, it is effective in practice. With these indices, layers in the decoding stage can easily be used to find original location information from the feature maps they received. However, unlike two other methods (bilinear and transpose convolutions, which were used in VGG(Visual Geometry Group) and FCN (Fully Convolutional Networks)), SegNet performs a non-linear upsampling of inputs ("unpooling" [6, 8–10]). Thus, the most important features and boundary delineation can be better reconstructed (Fig. 1).

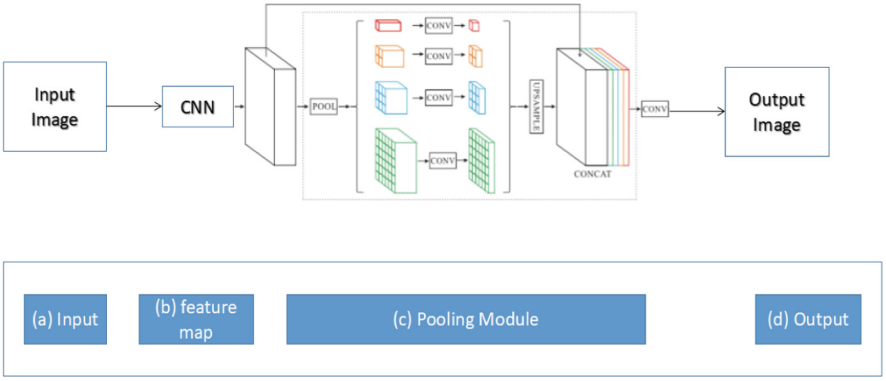


Fig. 1. Architecture of modified SegNet

To make an overall quantitative performance of the networks above, we select several measures: global accuracy (GA), class average accuracy (C), and mean

intersection over union (Mean IU). Let $\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}$ be the number of pixels of class i that are predicted to belong to class j (n stands for the number of different classes; in this paper, $n = 2$). Let T_i be the total number of classes, where $T_i = \sum_{j=1}^{n} p_{ij}$. We can compute the above accuracies as follows:

GA: the percentage of pixels in images that are classified correctly

$$\mathrm{GA} = \frac{\sum_{i=1}^{n} p_{ii}}{\sum_{i=1}^{n} T_i} \tag{1}$$

C: the mean of predictive accuracy of all classes

$$\mathrm{C} = \frac{1}{n} * \left(\sum\nolimits_{i=1}^{n} \frac{p_{ii}}{T_i} \right) \tag{2}$$

IU: the ratio of area of overlap to area of union

$$\text{Mean IU} = \frac{1}{n} * \left(\sum\nolimits_{i=1}^{n} \frac{p_{ii}}{T_i + \sum_{j=1}^{n} (p_{ji} - p_{ii})} \right) \tag{3}$$

3 Experimental Results

We conducted experiments on a workstation using Geforce GTX 1080 and 32 GB memory. Since 716 of the raw images have been labeled, we set $\mathrm{set}_1 = 200$ and $\mathrm{set}_2 = 716$ and the experiments respectively. We also compare U-Net and SegNet in later sections. The labeled images used in this experiment are shown as follows (Fig. 2):

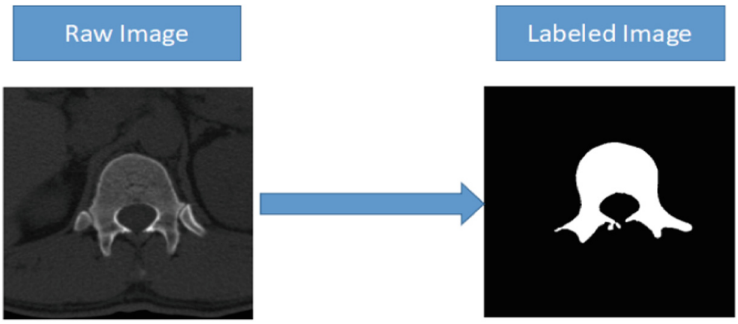


Fig. 2. A raw image and its labeled image

As we said in Sect. 2, we have images under different training data set sizes.

Table 1. Modified U-Net under set_1

Name	Epoch	Loss	GA	C	Mean IU
Modified U-Net	10	0.1444	93.12	74.21	67.47
	20	0.1506	93.08	77.25	72.44
	30	0.1552	92.97	73.45	71.33
	40	0.4734	89.4	64.22	54.12
	50	0.4933	89.06	63.64	53.15
	100	0.6214	87.59	61.57	50.62

We evaluate our segmentation results on spine CT images using three different accuracies for a detailed view of the efficiencies of our networks. From Table 1, we clearly see that the performance of U-Net is not bad, whatever the number of epochs. The basic structure and shape can be seen in Figs. 3 and 4. However, as the number of epochs increases, there is a great grain to loss function, which increases from 0.1444 to 0.6244, while the accuracies (of all three) decrease at the same time. The same thing happens in Table 2, which indicates a question of overfitting.

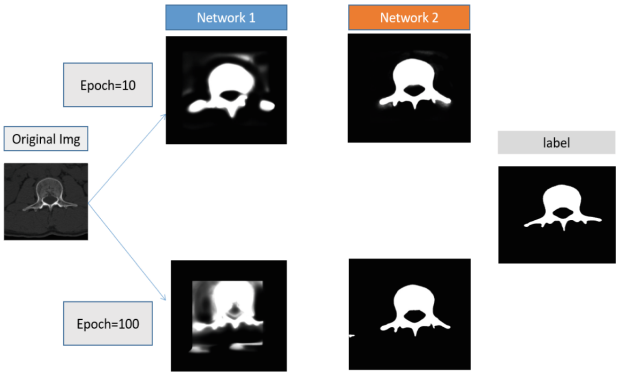


Fig. 3. Segmentation under modified U-Net

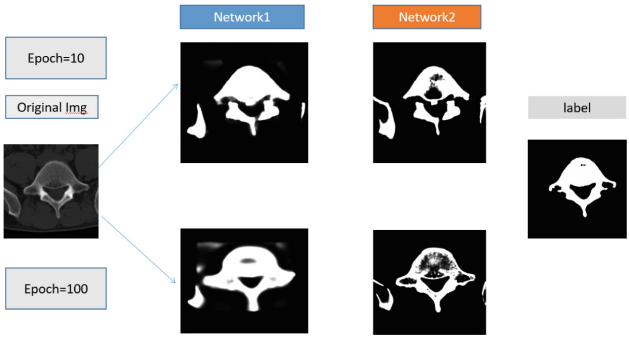


Fig. 4. Segmentation under modified SegNet

Table 2. Modified U-Net under set_2

Name	Epoch	Loss	GA	C	Mean IU
Modified U-Net	10	0.1329	94.33	71.66	65.95
	20	0.11122	94.43	71.16	70.92
	30	0.1259	94.8	72.63	69.81
	40	0.1240	94.893	73.24	52.6
	50	0.10063	94.95	76.01	51.63
	100	0.5216	89.86	60.01	49.1

Table 3. Modified U-Net under set_2

Name	Epoch	Loss	GA	C	Mean IU
Modified SegNet	10	0.1294	94.82	82.42	72.25
	20	0.1276	95.08	83.46	75.22
	30	0.1252	96.27	83.66	76.11
	40	0.1143	96.34	84.43	76.91
	50	0.1133	96.406	86.85	77.93
	100	0.1014	96.659	87.78	78.40

Table 4. Modified U-Net under set_2

Name	Epoch	Loss	GA	C	Mean IU
Modified SegNet	10	0.1379	94.79	83.99	73.04
	20	0.1222	95.11	85.03	74.01
	30	0.1158	96.24	85.23	76.94
	40	0.11056	96.32	86.57	77.07
	50	0.11298	96.501	87.42	78.12
	100	0.11294	96.459	88.05	78.29

From Tables 3 and 4, we can see SegNet has better accuracy but less loss when compared with U-Net, whatever the value of epochs. It can be concluded that SegNet is much more efficient than U-Net in experiments. Figures 3 and 4 show detailed segmentation processes for both U-Net and SegNet. It can be seen that with various values of epoch, the segmentation efficiency of U-Net decreases, which proves that overfitting occurs in U-Net. Conversely, SegNet does not have such an overfitting problem; the central intensity of the segmented image becomes higher with the change of epoch.

4 Conclusion

In this work, network-based medical image segmentation was demonstrated for spine CT. Two experiments were conducted using both U-Net and SegNet. We can conclude the following: (1) A U-Net-like network is preferred for segmentation applications

where the training set is not large. Otherwise, overfitting cannot be prevented. (2) Compared to U-Net, as demonstrated in Figs. 3 and 4, SegNet has better performance in segmentation, irrespective of the size of the training data. As for the reason for overfitting, considering that U-Net uses the training strategy of data augmentation by applying elastic deformations (by shifting, rotating, etc.) to the available training images, the training set used in every epoch is another version of the original images. With the increase in the number of training sets, the actual number of training sets U-Net used was much greater than that for SegNet. Too much training easily caused overfitting. Although Long et al. considered that data augmentation could not yield great improvements in the efficiency of models, it still plays an important role in particular fields.

Acknowledgments. This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No. 2017-0-018715, Development of AR-based Surgery Toolkit and Applications).

References

1. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: International Conference on Neural Information Processing Systems, pp. 1097–1105. Curran Associates Inc. (2012)
2. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
3. Dan, C.C., Giusti, A., Gambardella, L.M., et al.: Deep neural networks segment neuronal membranes in electron microscopy images. Adv. Neural. Inf. Process. Syst. **25**, 2852–2860 (2012)
4. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440 (2015)
5. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 234–241. Springer, Cham (2015)
6. Badrinarayanan, V., Kendall, A., Cipolla, R.: SegNet: a deep convolutional encoder-decoder architecture for image segmentation. arXiv preprint arXiv:1511.00561 (2015)
7. Kayalibay, B., Jensen, G., van der Smagt, P.: CNN-based segmentation of medical imaging data. arXiv preprint arXiv:1701.03056 (2017)
8. Noh, H., Hong, S., Han, B.: Learning deconvolution network for semantic segmentation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1520–1528 (2015)
9. Dumoulin, V., Visin, F.: A guide to convolution arithmetic for deep learning. arXiv preprint arXiv:1603.07285 (2016)
10. Ian, G., et al.: Deep Learning, vol. 1. MIT Press, Cambridge (2016)

Vehicular Fog Computing Based Traffic Information Delivery System to Support Connected Self-driving Vehicles in Intersection Environment

Joosang Youn(✉)

Department of Industrial ICT Engineering, Dong-Eui University,
176, Eomgwangno, Busan jin-gu, Busan 47340, South Korea
jsyoun@deu.ac.kr

Abstract. Recently, edge and fog computing server based various researches related on smart traffic control system have been studied to support enhanced traffic control service performance. In particular, fog computing server based novel ITS, such as vehicular fog computing, is being developed for traffic application of vehicle networks. Vehicles, regarded as smart devices with mobile, have the computational capability to collect useful traffic information about traffic state and road condition. Thus, in vehicular networks enabled to vehicular fog computing, through smart vehicular, fog server performing computing can be deployed at the edge of vehicular networks to collect, store, and compute traffic data related on vehicle state, road condition and traffic state. Here, vehicular fog computing server can provide a vehicle services to connected self-driving vehicle with the providing driving information, smart traffic control and road safety improvement. we propose a vehicle fog computing-based traffic information delivery system that can be used to deliver real-time and urgent-sensitive data, such as traffic data related on vehicle accident, in an intersection environment. The proposed system model is the functional architecture of traffic information delivery system based on vehicular fog computing server.

Keywords: Vehicular fog computing · Safety message · Vehicular networks

1 Introduction

Recently, fog computing technology is studied as potential solution to be able to be adapted at IoT environment [1, 2]. Fog computing is still in its early stage, with operational challenges, ranging from architecture to clear use cases to processing and computing issues and so on. In particular, fog computing based various researches related on connected self-driving vehicular fog computing have been studied to support enhanced service performance. Fog computing server based novel ITS technology, such as vehicular fog computing, is being developed for an vehicle application in vehicle networks. One potential application of fog computing is the integration of fog computing with conventional vehicular network to form the Internet of Vehicles (IoV) or vehicular fog computing [3]. Vehicular fog computing extends the fog

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 208–213, 2020.
https://doi.org/10.1007/978-981-13-9341-9_35

computing paradigm to conventional vehicular networks, which allows us to support more enhanced performance to driver, self-driving vehicle and vehicles application, by overcoming the limitations in conventional vehicular networks in terms of latency, location awareness, and real-time process through vehicular for computing. In vehicular networks enabled to vehicular fog computing, fog nodes performing computing can be deployed at the edge of vehicular networks to collet, store, and process traffic data related on vehicle from urban/highway areas. Vehicular fog computing can provide a vehicle-based services to connected self-driving vehicle with driving related information service, smart traffic control and road safety improvement. In this paper, we present vehicular fog computing architecture and a typical use case using vehicular fog computing. Then we propose potential solution which is latency-sensitive message broadcasting scheme based on vehicular fog computing to provide connected self-driving vehicles with the improved vehicular services connected self-driving vehicles. In particular, the proposed scheme can provide a method of rapidly disseminating latency-sensitive message such as accident notification message and safety message. The remainder of the paper is structured as follows. Section 2 summarizes the vehicular fog computing architecture and characteristics. Section 3 address proposed vehicular fog computing-based traffic Information delivery system. In Sect. 4, we show the performance. Finally, Sect. 5 conclude this paper.

2 Vehicular Fog Computing Architecture

A high-level system architecture of vehicular fog computing is shown in Fig. 1 [1, 3]. As shown in Fig. 1, vehicular fog computing consists of a vehicle, a fog node, and a cloud server. First, a vehicle collects data and plays a role of vehicle-level computing and decision through collected data. Fog nodes are located at Intersections or roads and collect various data from various vehicles and perform computation on the area covered by the fog nodes together with the convergence of collected data. Finally, the cloud server performs data analysis and city-level computing and determination of specific situations. In this paper, we propose traffic information delivery system based on the vehicular fog computing functions, addressed up. The proposed system can be used for an emergency/urgent traffic information delivery scenario related to safety vehicle applications such as vehicle accident situation and quickly transfers safety related messages to nearby vehicles.

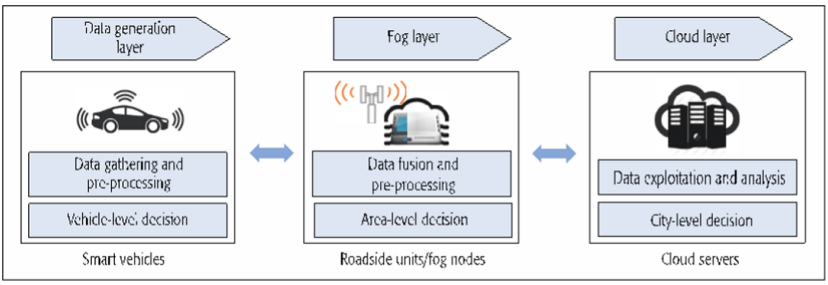


Fig. 1. System architecture of VFC [3]

3 Proposed Vehicular Fog Computing-Based Traffic Information Delivery System

In this paper, we propose a vehicle fog computing-based traffic information delivery system that can be used to deliver real-time and urgent-sensitive data, such as traffic data related on vehicle accident, in an intersection environment. Figure 2 shows the proposed functional architecture of traffic information delivery system based on vehicular fog computing server. Vehicle carries out the function of collecting vehicles information about the intersection situation, and the collected information is transmitted to the fog node. The fog node classifies the collected information into the status information to process the information contained in the information in the fog node or in the cloud server and it performs a judgment on whether or not there is a problem. If it is a traffic situation issue that requires local traffic issue or fast processing, it processes it first in the fog node. If it is an urban traffic issue, it processes the information collected after transmitting information to the cloud server. In this paper, vehicular fog computing based traffic information delivery system is proposed to provide connected self-driving vehicles with the improved vehicular computing and message delivery service. First at all, The assumed network is shown in Fig. 2. We use a vehicular fog computing server based traffic control & information delivery system as a use case to assume vehicular network using the vehicular fog computing. In the assumed network shown in Fig. 2, we assume that time-sensitive urgent traffic message related on road traffic congestion and vehicle accident in Intersection is generated and in this environment vehicular fog computing server performs procedure to broadcast the time-sensitive urgent traffic message to neighbor all vehicles including neighbor all vehicles connected to neighbor Intersection. For that, we propose traffic information delivery system based on vehicular fog computing server. a description of each element in the traffic information delivery system is shown in Fig. 2. A vehicular fog computing server monitors, manages and controls local traffic flow. Also, A vehicular fog computing server have communication range to be able to covers a region of intersection. If a connected self-driving vehicles are physically located within the communication range of a vehicular fog computing server, it can send and receive traffic information messages to and from the vehicular fog computing server. Especially, when there are a self-driving vehicles within the coverage of a vehicular fog computing server, it will frequently report traffic conditions, road conditions, accident information of its current location. Based on the traffic information received from a self-driving vehicles, the vehicular fog computing server can perform the following. While a vehicular fog computing server will monitor and control the local traffic flow of the intersection, if vehicle accidents occurs in the intersection, the vehicular fog computing server should quickly send accident messages to neighbor vehicles.

In order to perform this, in the proposed vehicular fog computing based traffic information system, the procedure is performed as shown in Fig. 3. Thus, the proposed scheme should operate safety message with real time and low latency. Also, the vehicular fog computing server will collect data and then pre-process and aggregate the received data and report such information to the cloud servers.

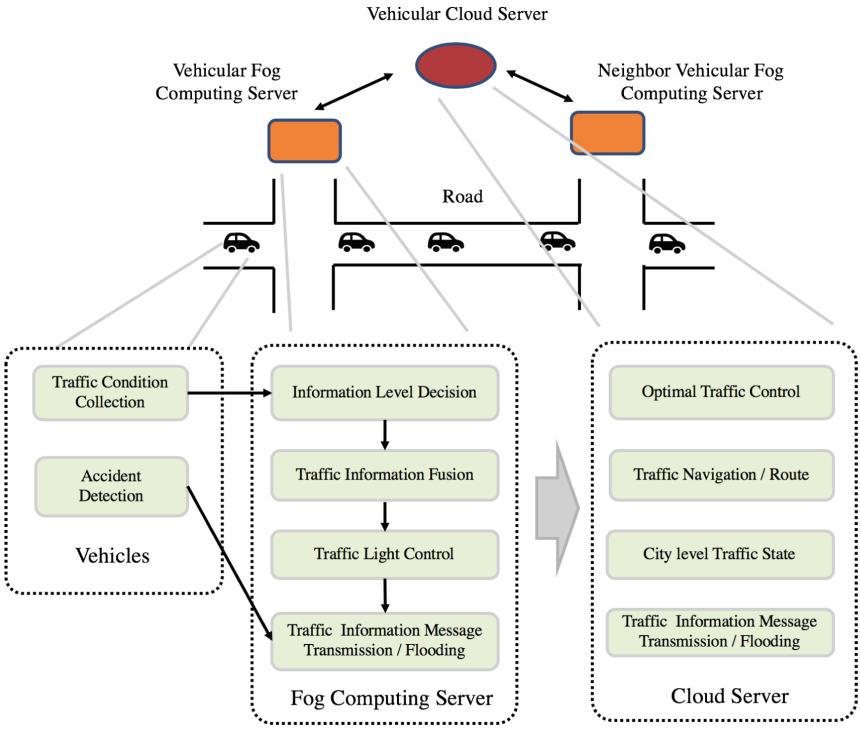


Fig. 2. The proposed VFC system functional architecture

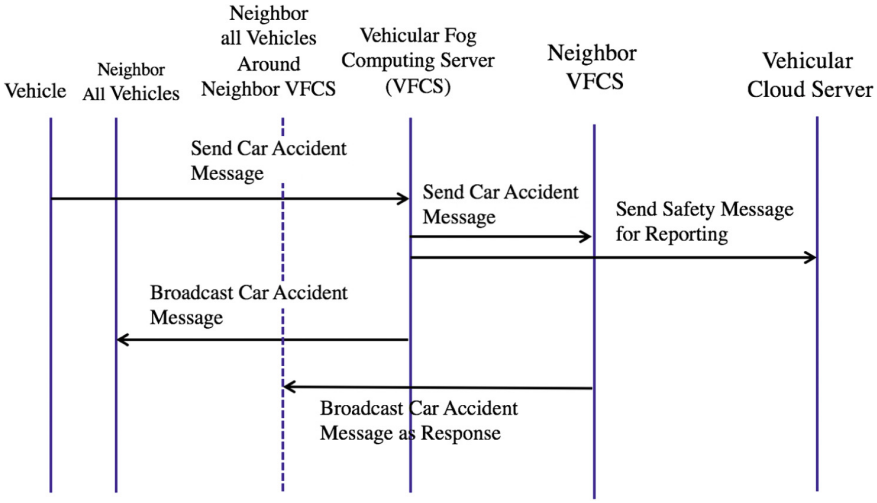


Fig. 3. The procedure of the proposed system

4 Performance Evaluation

In this section, in order to evaluate the performance of the proposed system, we use average message delivery time between vehicular, such as a key metric. In addition, we show the performance of the proposed system, compared with previously proposed cloud based general ITS via extensive simulations in NS (Network Simulator)-3. For these experiments, we consider simple network model shown in Fig. 1 for the simulation. In the simulation, two scenarios are considered as the network service model. One is evaluated in LTE-based VANET environment using vehicular fog computing based traffic information system. The other is evaluated in LTE-based VANET environment using Cloud based ITS. The simulation results are shown in Fig. 4. As shown in Fig. 4, the simulation result demonstrates that the proposed vehicular fog computing based traffic information system has lower delivery time than ITS based scheme and show that it is able to support the message delivery with short-term time between different vehicular fog computing sever. In the results, all flow traffic information is provided with stable transmission service in the VANET using the proposed vehicular fog computing based traffic information system. This shows that vehicular fog computing based traffic information system at the Intersection can distribute efficiently data transmission to neighbor vehicular fog computing based traffic information system at the neighbor Intersection.

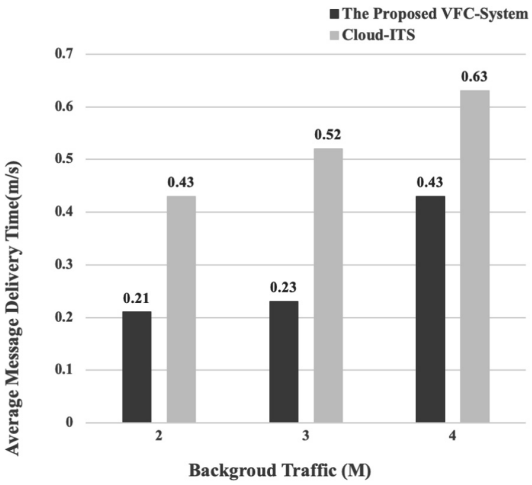


Fig. 4. The average message delivery time

5 Conclusion

In this paper, vehicular fog computing based traffic information delivery system is proposed to transmit urgent message for reporting vehicle accident to neighbor vehicles in vehicular fog computing based vehicular network environments. As explained up,

vehicular fog computing is the one of the best solution to support more enhanced vehicular performance for traffic message delivery service to driver, self-driving vehicle and vehicles application. Thus, the vehicular fog computing based traffic information delivery system can provide traffic control system with delivery safety delivery message with real-time delivery and with low latency service.

Acknowledgments. This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2017R1D1A1B03034689).

References

1. OpenFog Consortium: Openfog Architecture Overview (2017). https://www.openfogconsortium.org
2. Youn, J.: Latency-sensitive message broadcasting scheme based on vehicular fog computing for connected self-driving cars. In: Proceedings, International Conferences NGCIT 2018, ASTL 152, pp. 8–11. SERSC (2018)
3. Huang, C., Lu, R., Choo, K.: Vehicular fog computing: architecture, use case, and security and forensic challenges. IEEE Commun. Mag. **55**, 105–111 (2017)

A Research on Temperature Control System of Server Room Using IPMI

In-junOck and Min Choi(✉)

Department of Information and Communication Engineering,
Chung-Buk University, Cheongju-si, South Korea
dlswns9ll9@gmail.com, dr.choimin@gmail.com

Abstract. In this paper, we describe the design of remote monitoring program using IPMI driver included in BMC among the latest IT technologies. It is a program that controls the operation of turning the cooling system on and off according to the temperature by measuring the server room temperature. In this paper, we show the usability of the program, the order of kernel compilation, the directionality, necessary system, some source code, and Ganglia monitoring screen.

Keywords: Remote control · IPMI driver · Computer system

1 Introduction

In entering the fourth industrial revolution, many companies and developers are developing numerous convenient systems and IoT devices. These devices can connect to the Internet from anywhere and are not restricted by the location and age [1]. A few years later, it is expected that many systems in daily life will be open to the era of automation or artificial intelligence.

However, it is not easy to manage servers, as various IoT products must communicate with each other and be connected to the Internet. In general, a good environment for users to use is felt convenient when the speed is fast, there is no inconvenience in using, and the user can quickly and easily obtain the desired result. As a condition for becoming such a system, it is the server that thinks it is important among many things. Since the server is the user's communication means, I think that good results will be obtained with the appropriate server.

As a condition for becoming a good server, I think that there are various things such as high-end (high-spec) server computer, computer cooling system, program algorithm, computer environment etc.

The most sensitive part of the various conditions is the temperature. As the temperature of the CPU increases, you can see that the performance is lower than when it is at the proper temperature. Therefore, companies and individual developers also strive to build a good cooling system to maintain the CPU temperature. Users and developers and businesses are familiar with the importance of servers.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 214–218, 2020.
https://doi.org/10.1007/978-981-13-9341-9_37

The company strives to optimize the server environment to provide a better server. For developers, efforts to build a good environment and system, for these reasons, management costs of servers in enterprises tend to increase annually. I believe that this system has been able to reduce costs when used by companies that need to manage servers [2].

While there are various IPMI monitoring systems, we plan to construct a system using Ganglia capable of scalable distributed monitoring for high-performance computing systems such as clusters and grids.

2 Background

2.1 IPMI Background Knowledge

There are kernel, IPMI, BMC, kernel compilation, etc. as background knowledge necessary for the system to be implemented in this paper. First, the kernel manages memory, processes, etc. in the operating system, and manages resources so that the system can operate normally. Compiling refers to the process of making user-created source code into an executable file using a compiler. Since it is an open source operating system which released the source code of Linux, it is possible for users to directly change and the process of fixing is called kernel compilation [3]. The following Fig. 1 shows the internal structure of the Linux kernel.

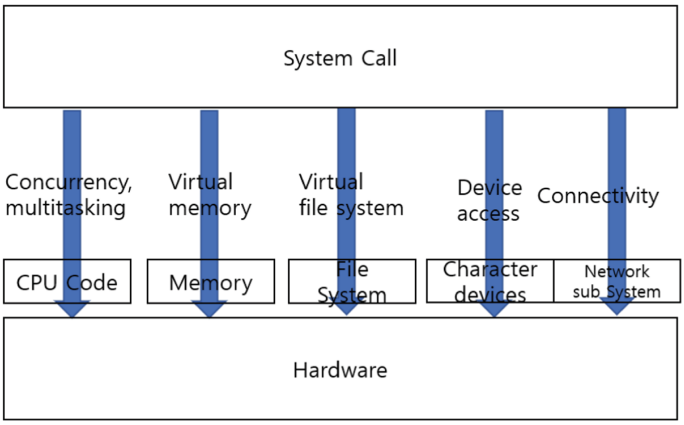


Fig. 1. Simplified kernel internal structure

Next, IPMI is a Device Driver that allows you to check and control the state of the server regardless of its location as a management interface for server management. IPMI is a driver that is built into a BMC capable board and IPMI can not be used if it is a board that does not support BMC. Next Fig. 2 is a block diagram of IPMI.

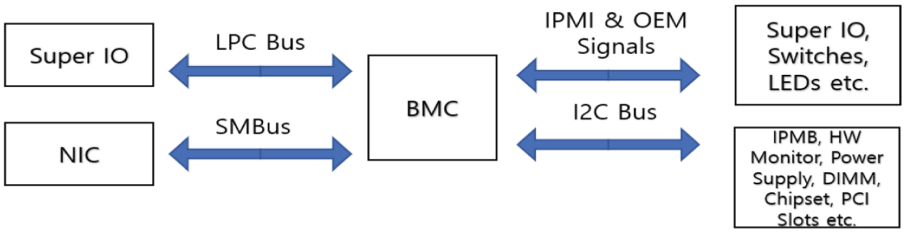


Fig. 2. IPMI block diagram

The IPMI 1.0 standard was enacted in a forum where more than dozens of companies have participated, mainly IntelTM and HP [1]. The IPMI information is exchanged via the BMC of the hardware component of the IPMI specification. It is a controller that manages the board called BMC and Baseboard Management Controller. The BMC's feature is designed to operate using the standby power supply even when the system is turned off.

Moreover, it can manage the system of only BMC. In order to manage the local node, if there is an IPMI driver and an application supporting him, it is done. Figure 3 shows a brief overview of the BMC system management architecture.

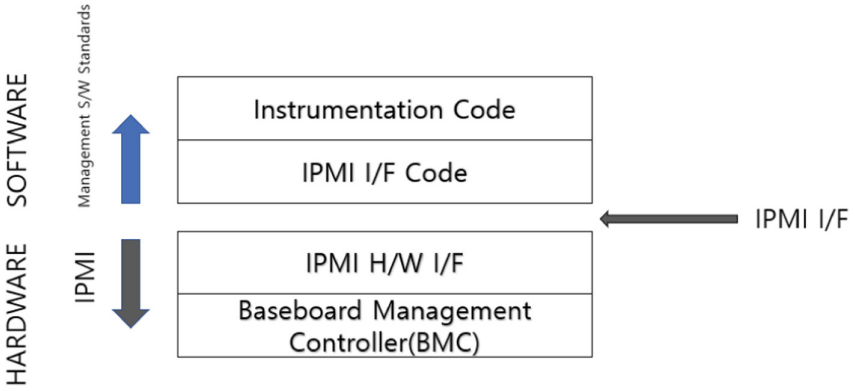


Fig. 3. Simplified BMC system management structure Source: www.intel.com

2.2 Performance Changes According to CPU Temperature

A protection function is added in the latest Intel CPU. It is a function to control the CPU clock by temperature conversion of the CPU and lower the temperature. In this way, it is possible to know the fact that the clock rotation number of the CPU influences the temperature (Fig. 4).

The CPU can see that performance changes are intense depending on the temperature. When the temperature rises, the symptoms first displayed are too much

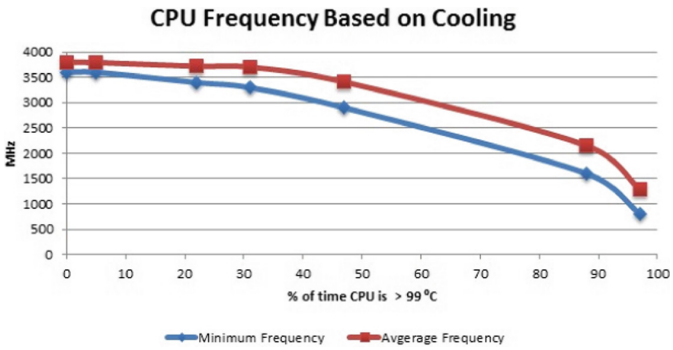


Fig. 4. CPU frequency based on cooling Source: http://jisan.berrytour.com/164 [4]

impact on the server due to performance degradation. If the temperature is maintained at high temperature for several days, the CPU is stopped. This phenomenon is said to be down [5].

3 Experiment and Consideration

We tested the CPU clock speed according to the CPU temperature. The following figure shows the results of evaluating the performance according to the CPU temperature. The horizontal axis shows temperature and the vertical axis shows CPU clock speed with temperature. We conducted a Stress Test to increase the temperature of the CPU (Fig. 5).

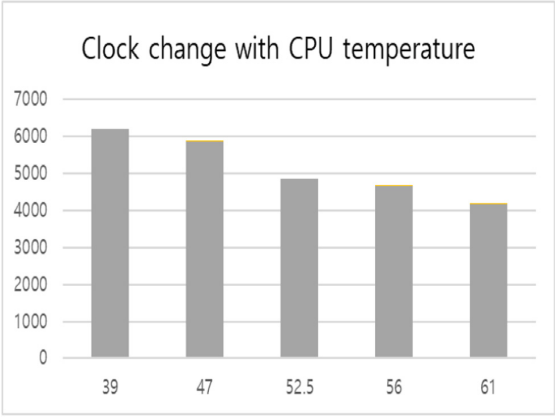


Fig. 5. Clock change with CPU temperature

For the program used for Test, Core Temp, which is public software, conducted Test in C language using assembly. The performance of the CPU rises by 5 to 6 degrees, the rotation speed of the core clock has a difference of at least 300 and a maximum of 1000. At first glance there seems to be no big difference If the server handles a lot of quantity at once, the story will be different. Also, if you go over 50 degrees, the number of clocks will drop to 1000 at once. From these results, it can be seen that the CPU maintains a certain number of revolutions at a certain temperature, and when the CPU exceeds the limit, the performance suddenly deteriorates.

4 Conclusion

As we enter the fourth industrial revolution, many IoT devices have appeared, and we enjoy the convenience that these devices provide. Developers who do not give convenience to users only have to build a system that can manage easily and conveniently by using useful functions.

In this paper, the direction using IPMI is presented to construct such a system. IPMI is superior to software for diagnosing servers reasons, you can manage systems in multiple physical locations. That is, it is firmware running on the motherboard of the system, not dependent on the operating system of the system.

As can be seen in this paper, the CPU is affected by temperature. In the case of my computer, the clock performance decreased by about 300 as the number 5 increased. At the present time, the clock performance declined by about 1000 degrees. IPMI can remotely monitor the temperature, join the program, monitor, and plan to operate the cooling system in stages.

Acknowledgement. This research was supported by strategic research project of NRF-2017R1E1A1A01075128 by National Research Foundation.

References

1. Kim, S.-W., Kim, D.W., Kim, S.-W.: Design and implementation of hardware resource monitoring system based on IPMI. In: Conference on the Institute of Electronics Engineers of Korea, pp. 433–434 (2009)
2. Kim, H., Na, J.-H., Park, S.-H., Kwak, S.: A Wifi Smart power outlet for remote monitoring and control of power consumption. J. Korea Multimed. Soc. **17**(2), 160–169 (2014)
3. http://www.linux4d.net/lvsp_org/community/source/3rd/html/g.html
4. http://jisan.berrytour.com/164
5. Lim, S.-S., Son, M.-K., Sung, W.-K., Kim, B.-H., Cho, D.-U.: Comparison and analysis of CPU temperature according to the program by used. In: Proceedings of Symposium of the Korean Institute of communications and Information Sciences, pp. 978–979 (2011)

Software Test Effort Estimation Based on Source Code Change History and Defect Information

Jongse Won[1] and Yeong-Seok Seo[2(✉)]

[1] Mobile Division, Samsung Electronics, 129 Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do 16677, Republic of Korea
[2] Department of Computer Engineering, Yeungnam University, 280 Daehak-Ro, Gyeongsan, Gyeongbuk 38541, Republic of Korea
ysseo@yu.ac.kr

Abstract. Software test effort estimation has always been an important activity for meeting testing deadlines, relocating resources, and reducing testing cost. However, in practice, estimating test effort is still a major challenge because it is difficult to quantify and collect factors that affect accurate test effort. In this study, we propose a new test effort estimation model by analyzing the relationship between the number of defects collected through defect tracking tools and the source code changes collected through configuration management tools during the software development period. Experiments are performed to validate the proposed model using real industrial software project data. The results indicate that the proposed model achieves high estimation accuracy.

Keywords: Software test effort estimation · Software defect · Configuration management · Software project management · Correlation

1 Introduction

Once the software implementation phase is complete in the software development phase, the software goes through the testing phase to improve quality and find defects. Software testing plays an important role in software quality assurance by executing a program with the intent of finding the software bugs, which has a significant effect on the overall quality of the final software product. Note that software testing is a time and resource consuming process. In general, testing effort accounts for 40%–50% of total resources, 30% of total effort and 50%–60% of the total cost of software development [1].

In order to perform a successful testing phase, test effort estimation is a crucial activity to plan the testing schedule accurately and efficiently to meet deadlines. Generally, reliable estimation is in the list of top ten factors for the success of software development projects [2, 3]. However, the test effort estimation is very difficult because uncertainty is unavoidable when predicting the future; it is also impossible to determine all factors that affect software testing and difficult to quantify and collect most factors (the total effort required for testing depends on various factors including human factors,

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 219–227, 2020.
https://doi.org/10.1007/978-981-13-9341-9_38

process issues, testing techniques, tools, and characteristics of the software development artifacts). In addition, estimating the effort required for a software project is more difficult as the project size increases. Inaccurate effort estimates may lead to poor software quality, discouraged project members, and customer dissatisfaction [3].

In this paper, we propose a new test effort estimation model using the source code changes collected through configuration management tools and the defects collected through defect tracking tools, with the technical and environmental factors. The basic idea is that the source code change during the software development is closely related to defects [4, 5], and this relation directly affects to test effort. Thus, we investigate the relationship between total changed lines of code in the committed file classified by types and the number of defects, and apply it to build the model. In our experiment, we evaluate the accuracy of the proposed model with various criteria using real industrial project data.

The remainder of this paper is organized as follows: Sect. 2 introduces existing studies for test effort estimates. Sections 3 describes the proposed test effort estimation model. Section 4 shows the experimental design and analysis, and finally Sect. 5 concludes the paper.

2 Related Work

Test effort estimation is an approximate calculation of effort related to the testing of a specified software development project in a particular environment using defined methods, techniques, and tools. There are several models that have been developed to estimate the test effort.

De Almeida et al. [6] describe a test effort estimation method based on the use case. They consider the technical and environmental factors as well as actor ranking. Srivastava et al. [7] propose a test effort estimation model that exploits the advantages of cuckoo search for effort optimization. Islam et al. [8] develop a novel tool for a test estimation technique that provides the time and cost using use cases and functions. Jayakumar et al. [9] present an overview of software test estimation techniques surveyed: Judgement and rule of thumb-based techniques, Analogy and work breakdown techniques, Factor and weight-based techniques, Functional size-based techniques, and Training-based techniques. They also provide some of the challenges that need to be overcome.

Although the existing studies are helpful to perform estimating test effort, they usually focus on the static factors (e.g., size and complexity). This is a crucial limitation in that software keeps changing until finishing the software development. This is especially important for test effort estimated when software development is almost complete. In order to solve the above issue, we consider the configuration management and defect tracking tools that include dynamic environment during the software development.

3 Test Effort Estimation Model

The overall approach is described in Fig. 1. In the following subsections, we present each step in more detail.

Fig. 1. Overall approach

3.1 Step 1: Calculating the Basic Test Effort Based on Software Size

Software consists of components (parts of the software) and their interoperations. In this step, the basic test effort for each component is calculated using its size.

The easiest way to measure the software size is by the number of source lines of code (SLOC) written. SLOC is a metric to measure the software size by counting the number of lines of code; the software size is mostly proportional to SLOC. After obtaining SLOC for each component, we can generate the basic test effort by multiplying the SLOC with the basic effort per SLOC (unit size). This basic effort per SLOC can be derived from the previous software project data obtained from within the software organization.

3.2 Step 2: Analyzing the Configuration Management Information

Software developers implement each part of software program solutions with checking whether there are any potential defects. However, they have different behavior styles regarding implementation, and the developer's behavior can affect the test effort required later. Thus, we identify the behaviors as the factors for estimating test effort, using the configuration management system. Table 1 shows the factors from configuration management information.

SLOC in the Committed File Classified by Type. When developers perform the commit action, the source code is saved into a repository and synchronized with the old source code committed by other developers. If the added software source code is large, the probability of defects is also high, which is proportional to its size. The classification of these source code modifications is as follows: (1) the number of SLOC added (adding new source code creates new defects related to the additional functions), (2) the number of SLOC deleted (deleting the existing source code reduces the probability of

Table 1. Factors obtained from configuration management information

	Configuration management information	Types
1	SLOC in the committed file classified by type	Added
2		Deleted
3		Updated
4		Total
5	Number of committed files classified by size	Small (SLOC < 100)
6		Medium (100 ≤ SLOC < 500)
7		Large (500 ≤ SLOC < 1,000)
8		Very large (SLOC > 1,000)
9		Total

defects), (3) the number of SLOC updated (updating the existing source code removes the defects), and (4) the total number of SLOC changed (this is the sum of SLOC of all types of codes changed).

Number of Committed Files Classified by Size. Some developer can commit large source code at one time but another developer commits small source code frequently. Thus, the number of committed files based on their sizes can be also important factor along with SLOC in the committed file classified by type. We categorize this size as (1) small (SLOC < 100), (2) medium (100 ≤ SLOC < 500), (3) large (500 ≤ SLOC < 1,000), (4) very large (SLOC > 1,000), and (5) total (total number of commit actions).

3.3 Step 3: Deriving Relative Weights Using the Configuration Management Information and the Number of Defects

In order to increase the accuracy of the entire test effort, we need efficient test effort allocation for each component. In this step, relative weighting is applied based on the configuration management information and the number of defects.

Weighting Depending on the Importance of the Factors Obtained from Configuration Management Information. In the previous subsection, nine configuration management factors are obtained. Because the defect occurrence by each factor is different, relative weighting is applied in the proposed estimation model. In this study, to obtain the weights, the Pearson correlation coefficient [10] is used between the number of defects and the configuration management factors.

Weighting Depending on the Size of the Configuration Management Information on Each Component. If the size of the configuration management information is larger, the test effort is higher. Thus, this is considered in the proposed estimation model, and relative weighting depending on the size for each component is performed. In this study, the weight is calculated by z-score [10] $((x_i - \mu)/\sigma$, where x_i is the size of configuration management information for i-th component, μ is the average size of the configuration management information for each component, and σ is the standard

deviation of the configuration management information size for each component). The ranges are set as from −1 to 1 for relative standardization.

3.4 Step 4: Identifying Technical and Environmental Factors

During the software testing period, the testing activity is influenced by various technical and environmental factors as well as the tester's performance [9]. There are many technical and environmental factors such as the availability of test tools, test plans, and test cases. In this study, the commonly used factors are applied in our proposed estimation model, which are illustrated in Table 2 along with their assigned weights [6, 11, 12].

Table 2. Technical and environmental factors

Complexity factor (CF)	Description	Assigned weight
1	Test tools	5
2	Documented inputs	5
3	Development environment	2
4	Test Environment	3
5	Test-ware reuse	3
6	Distributed system	4
7	Performance objectives	2
8	Security features	4
9	Complex interfacing	5

Each factor receives a value, which represents degree of importance (DI) to the software testing. The range is from zero to four (Harmless = 0, Needless = 1, Desirable = 2, Necessary = 3, Essential = 4). The multiplication between the assigned weight and DI is named as *CF*, and Technical and Environmental Factor (*TEF*) is calculated by the sum of *CF* from all factors. The equation is as follows [6, 11, 12]:

$$TEF = 0.65 + 0.01 * \sum CF_i. \tag{1}$$

3.5 Step 5: Generating the Total Test Effort Estimate

We obtain the total test effort (*TTE*) from the values in the previous steps. This is expressed as follows:

$$\boldsymbol{TTE} = \left[\sum_{i=1}^{n} \left\{ \sum_{j=1}^{m} (BE)_j * (f_i) * \left(1 + g_{ij}\right) \right\} \right] * TEF. \tag{2}$$

where *BE* is the basic test effort (in Step 1), f_i is the *i*-th factor's relative weight depending on the size of the correlation coefficient between the configuration

management information and the number of defects (in Step 2 and 3), g_{ij} is the relative weight by the size of the configuration management information in i-th factor and j-th component (in Step 3), and *TEF* is the weight related to the technical and environmental factors (in Step 4). Because the range of g_{ij} is from −1 to 1, 1 is added for g_{ij} to avoid having a negative value for the test effort estimate on each component.

4 Experiment

The proposed model is validated using the real software project data collected from one of the conglomerate companies in Korea. The project used for this experiment is a mobile application implementation in the year 2012; the implementation period of the application was eight months. It was released 19 times officially and 25 developers were committed to the project. In the project, the configuration management tool used was Perforce and defect tracking tool used was ClearQuest. Experimental design is as shown in Fig. 2. In our experiment, we measure the estimation accuracy by comparing the actual test effort with the estimated test effort from the proposed model.

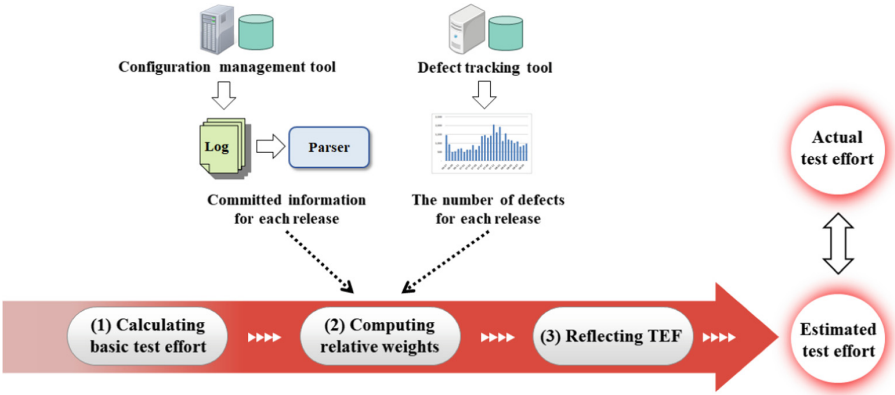


Fig. 2. Experimental design

As described in Subsects. 3.2 and 3.3, the relative weights are derived using the configuration management information and the number of defects. One of the weights obtained from the correlation coefficients among 19 releases are presented in Table 3.

As described in Subsect. 3.4, *TEF* is derived based on the CF in Table 2. In order to calculate CF, by surveying the one team member participated in the project, DI is assigned as follows: Test tools = 4, Documented inputs = 4, Development environment = 4, Test environment = 4, Test-ware reuse = 4, Distributed system = 3, Performance = 4, Security features = 3, Complex interfacing = 4.

In order to measure the accuracy of the proposed model, as shown in Table 4, the most common criteria for evaluating the accuracy of effort estimation are used in our experiment [13].

Table 3. Results of the relative weights in a release

	Configuration management information		Defect information	Correlation coefficients
1	SLOC in the committed file classified by type	Added	The number of defects in a release	0.698
2		Deleted		0.799
3		Updated		0.737
4		Total		0.753
5	Number of committed files classified by size	Small (SLOC < 100)		0.771
6		Medium (100 ≤ SLOC < 500)		0.751
7		Large (500 ≤ SLOC < 1,000)		0.459
8		Very large (SLOC > 1,000)		0.622
9		Total		0.754

Table 4. Evaluation criteria

Evaluation criteria	Description
MMRE (The Mean Magnitude of Relative Error): the average MRE_i	$= \frac{1}{n}\sum_{i=1}^{n} \underbrace{\frac{\lvert ActualEffort_i - EstimateEffort_i \rvert}{ActualEffort_i}}_{MRE_i}$
Pred(0.25)	The % of the data with MRE ≤ 0.25
Pred(0.30)	The % of the data with MRE ≤ 0.30

We compare the estimated test effort for all 19 releases with the actual test effort, and the experimental results are presented in Table 5. MMRE is 0.2326, and Pred(0.25) is 68% and Pred(0.30) is 100%. Note that MMRE ≤ 0.25 is the standard criterion to be considered as an accurate and acceptable level for an estimation model [14]. Pred(0.30) is 100%, which indicates that 100% of the releases fall within 30% error range.

Table 5. Results of test effort estimation (MRE)

Release ID	Estimated effort	Actual effort	MRE
1	697.6	585	0.1925
2	695.4	589	0.1806
3	754.7	589	0.2813
4	712.5	590	0.2076
5	723.4	593	0.2199
6	730	594	0.229
7	719.9	595	0.2099

(*continued*)

Table 5. (*continued*)

Release ID	Estimated effort	Actual effort	MRE
8	759.5	599	0.2679
9	768.7	602	0.2769
10	756.4	603	0.2544
11	734.8	604	0.2166
12	755.9	606	0.2474
13	746.7	607	0.2301
14	742.4	609	0.219
15	770.3	609	0.2649
16	743.9	609	0.2215
17	738.8	609	0.2131
18	751.7	610	0.2323
19	766.3	611	0.2542
MMRE			**0.2326**
Pred(0.25)			**68%**
Pred(0.30)			**100%**

5 Conclusion

It is important to estimate test effort correctly to maximize quality and minimize cost of software development. In this paper, we propose a novel test effort estimation model using the committed information collected through configuration management tools and the defects collected through defect tracking tools, with the technical and environmental factors. The performance of the proposed model is evaluated by the accuracy measures with the real software project data collected from one of the conglomerate companies in Korea.

Acknowledgment. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2017R1C1B5018295).

References

1. Kumar, D., Mishra, K.K.: The impacts of test automation on software's cost, quality and time to market. Procedia Comput. Sci. **79**, 8–15 (2016)
2. Abran, A.: Software Project Estimation: The Fundamentals for Providing High Quality Information to Decision Makers. Wiley-IEEE Computer Society Press, Los Alamitos (2015)
3. Nasir, M.H.N., Sahibuddin, S.: Critical success factors for software projects: a comparative study. Sci. Res. Essays **6**, 2174–2186 (2011)
4. He, Y., Zhu, X., Wang, G., Sun, H., Wang, Y.: Predicting bugs in software code changes using isolation forest. In: 2017 IEEE International Conference on Software Quality, Reliability and Security, pp. 296–305 (2017)

5. Kim, S., Whitehead, E.J., Zhang, Y.: Classifying software changes: clean or buggy? IEEE Trans. Softw. Eng. **34**, 181–196 (2008)
6. de Almeida, E.R.C., de Abreu, B.T., Moraes, R.: An alternative approach to test effort estimation based on use cases. In: 2009 International Conference on Software Testing Verification and Validation, pp. 279–288 (2009)
7. Srivastava, P.R., Varshney, A., Nama, P., Yang, X.S.: Software test effort estimation: a model based on cuckoo search. Int. J. Bio-Inspired Comput. **4**, 278–285 (2012)
8. Islam, S., Pathik, B.B., Khan, M.H., Habib, M.M.: A novel tool for reducing time and cost at software test estimation: an use cases and functions based approach. In: 2014 IEEE International Conference on Industrial Engineering and Engineering Management, pp. 312–316 (2014)
9. Jayakumar, K.R., Abran, A.: A survey of software test estimation techniques. J. Softw. Eng. Appl. **6**, 4–52 (2013)
10. Devore, J.L.: Probability and Statistics for Engineering and the Sciences. Cengage Learning, Boston (2015)
11. Nageswaran, S.: Test effort estimation using use case points. Qual. Week **2001**, 1–6 (2001)
12. Sharma, A., Kushwaha, D.S.: An empirical approach for early estimation of software testing effort using SRS document. CSI Trans. ICT **1**, 51–66 (2013)
13. Wu, D., Li, J., Bao, C.: Case-based reasoning with optimized weight derived by particle swarm optimization for software effort estimation. Soft. Comput. **22**, 5299–5310 (2018)
14. Kitchenham, B.A., Pickard, L.M., MacDonell, S.G., Shepperd, M.J.: What accuracy statistics really measure. IEE Proc. Softw. **148**, 81–85 (2001)

Mobile Health Monitoring System Including Biofeedback Training Through Analysis of PPG and Respiratory Pattern Change

Daechang Kim[1], Jaeyong Kim[1], Jaehoon Jeong[2], and Sungmin Kim[2(✉)]

[1] Department of Medical Biotechnology, Dongguk University-Bio Medi Campus, Seoul, Gyeonggi-do 10326, Korea
kimdaechangl0@gmail.com, ballentain68ll@gmail.com
[2] Department of Medical Industry, Dongguk University-Seoul, Seoul 04620, Korea
bfsl.jeong@gmail.com, sungmin2009@gmail.com

Abstract. Free radical oxygen generated by stress and oxygen deficiency causes problems such as heart disease and stroke. We propose mobile health monitoring system to solve this problem. Based on biofeedback technology used in psychiatry, meaningful parameters were identified by analyzing the autonomic nervous system and breathing patterns that change during deep breathing. As a result, PP-interval was increased ($p < 0.05$) during deep breathing and the average expiration value of the subject was unconsciously increased ($p < 0.05$). The deep breathing has been shown to be able to prevent stress and oxygen deficiency. After that, we developed an effective mobile health monitoring system using parameter that have significant result value. If personal health monitoring system that continuously induces deep breathing through the method of this study is developed, it can be used as a way to continuously manage and improve individual health.

Keywords: Ubiquitous system · Biofeedback · Breathing pattern · PP-interval

1 Introduction

The ubiquitous system, which can connect people, computers and objects together to provide optimized information processing and information, is utilized in various industries and provides flexible services. Among them, ubiquitous system is used in diabetes, stress and hypertension management in medical field. In addition, it plays a role of providing personal a biological signal to users by utilizing a simple biological signal monitoring system [1, 2]. However, there are few methods available to promote health using mobile.

There is a way to improve health by utilizing biological signal. The mental health department uses biofeedback, a health promotion training method that uses biological signal information. The biofeedback training is to make the patient aware of the change in the biological signal information of the actual patient and lead to a change in the consciously physiological activity itself [3]. Biofeedback training is aimed at activation

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 228–233, 2020.
https://doi.org/10.1007/978-981-13-9341-9_39

of parasympathetic nervous system. Of all the training methods, the deep breathing method is used as an initial training. Deep breathing activates parasympathetic nervous system as it supplies a sufficient amount of oxygen and help the cell.

The importance of deep breathing is related to the free radical oxygen produced in the human body. Free radical oxygen can be generated by normal metabolic processes, but cells directly generate free radical oxygen due to lack of oxygen and stress. Large amounts of free radicals oxygen damage cell membranes, chromosomes and proteins and resulting in arteriosclerosis, heart disease and stroke [4–6]. In other words, it will be important to present measures to proactively recognize and resolve stress, lack of oxygen, etc. which can increase large amounts of free radicals oxygen.

Photoplethysmography (PPG), which is widely used in mobile devices, can analyze pulse waves using change in blood volume of microvascular layer [7]. The frequency domain of heart rate variability, which is a change in Peak to Peak (PP-interval), is generally divided into a Low frequency band (LF) and a High frequency band (HF), which represent sympathetic and parasympathetic nerves, respectively. When the HF increases, the PP-interval increases. By analyzing the PP-interval, the change of the autonomic nervous system can be confirmed indirectly [8]. In other words, the change of the PP-interval value can confirm stress.

Biofeedback has the disadvantage of having to visit the place of training because it has to be continuously trained for 4 weeks. To solve this problem, we propose Framework of mobile health monitoring system. In this study, PPG signal was used to facilitate the use in health promotion system. The autonomic nervous system changes were observed by comparing PPG data of resting state and PPG data of breathing training of biofeedback. This value will be used to detect the stress and to confirm the effectiveness of the training in the framework. Next, physiological changes were observed by changing respiration. This shows physiological change that are unconsciously changed after deep breathing. We provided parameters to effectively use the proposed monitoring system. If using the parameters and system presented in this study, it is expected that it will help to develop ubiquitous system that can substantially improve individual's health while overcoming the shortcomings of biofeedback training.

2 Method

Experiment was performed under the same environment conditions where the external stimulus was controlled. PPG was measured using BioPack M36 and SS4LA pulse sensor. Breathing volume were measured by using a SS5LB respiration sensor. The subjects were 10 without cardiovascular disease such as arrhythmia and myocardial infarction. Subjects were not informed that the purpose of the detailed experiment. This prevented the subject from acting consciously. After all the tests were completed, the subjects were briefed on the purpose of the experiment and we received an informed consent.

Signal analysis was performed using the MATLAB R2015b version. First, we obtained data of PP-interval, expiration and inspiratory value in the three step of Fig. 1. Through the acquired information, we compared with the respiration size of before breathing zone and after the deep breathing zone for confirming unconscious physiological changes of subjects. Additionally, to determine if deep breathing affects the parasympathetic nerves, we measured the PP-interval by extracting the respiration time of the deep breathing zone.

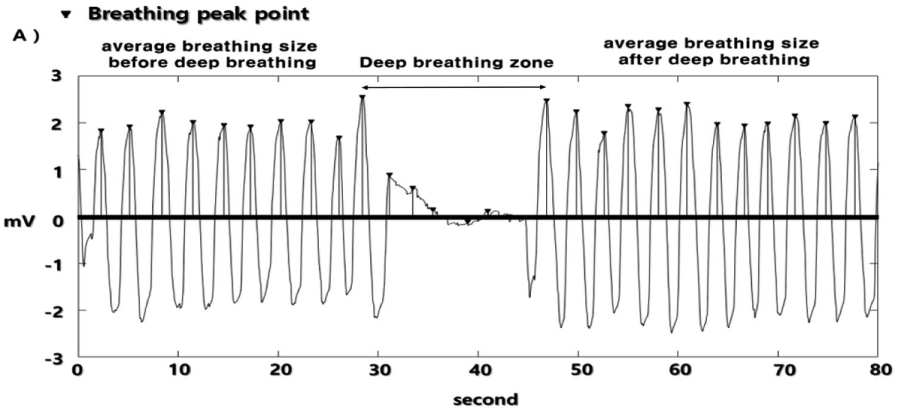


Fig. 1. Biofitback Deep breathing training change graphs (a) PPG changes during deep breathing, (b) changes in breathing size before and after deep breathing

3 Result

Table 1 shows the parameters that can be used in the monitoring system. In order to observe the physiological changes, the change in the volume of the maximum expiration and the maximum inspiratory before and after deep breathing were calculated. After deep breathing in all subjects, the volume of the exhalation breathing increased like Table 1 ($P < 0.05$). The increased volume of exhalation means that more air escapes from the lungs. This can be understood as an increase in carbon dioxide emissions. In the case of inspiratory volume, the change in the subjects were not the same. However, the total volume increased slightly or had similar values. Also, at deep breathing, the value of the RR-interval increases during deeper breathing than the average RR-interval. This can be thought of as the active parasympathetic nervous during deep breathing, making the heart rate low and relaxing. The values of Table 1 were found to be $P < 0.05$, which is the average RR-interval (0.8979) and the deep breathing RR-interval (0.9437) through t-test (Fig. 2).

Table 1. Parameters available in the monitoring system

Subject	Average breathing (s)	PP-interval in average breathing (s)	Expiration average expiration value before deep breathing (mV)	Deep breathing (s)	PP-interval in deep breathing (s)	Expiration average expiration value after deep breathing (mV)
1	5.26	0.9674	1.8644	15.32	1.03	2.2106
2	2.88	0.9816	3.2955	14.94	1.03	6.7481
3	3.26	0.9176	1.6494	20.31	0.9472	2.2649
4	3.35	0.6627	1.3926	18.39	0.7305	2.4374
5	6.63	1.2038	1.6632	18.44	1.2817	1.7348
6	3.31	0.928	2.9508	20.32	0.9529	4.5355
7	3.82	0.9183	2.9099	17.48	0.9738	3.4527
8	2.5	0.8056	2.343	19.57	0.8797	2.4825
9	3.97	0.8117	1.5433	17.2	0.8236	1.6455
10	3.21	0.7832	1.7398	17.37	0.7876	2.386

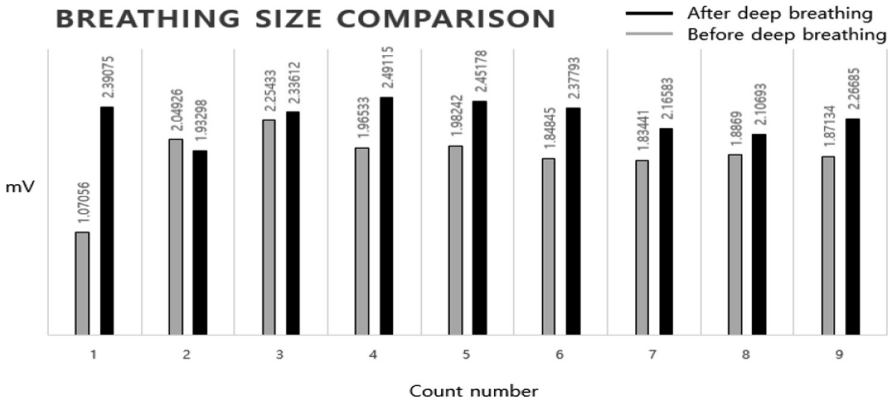


Fig. 2. Expiration breathing pattern analysis by comparison before deep breathing and after deep breathing

After deep breathing, the volume of breathing increased during 9 breaths of the subject. This suggests that the parasympathetic nervous system activated by deep breathing consistently induce deep breathing patterns. The number of increased breaths was applied as a parameter to determine proper training results. The monitoring system as shown in Fig. 3 is presented through these synthesized features.

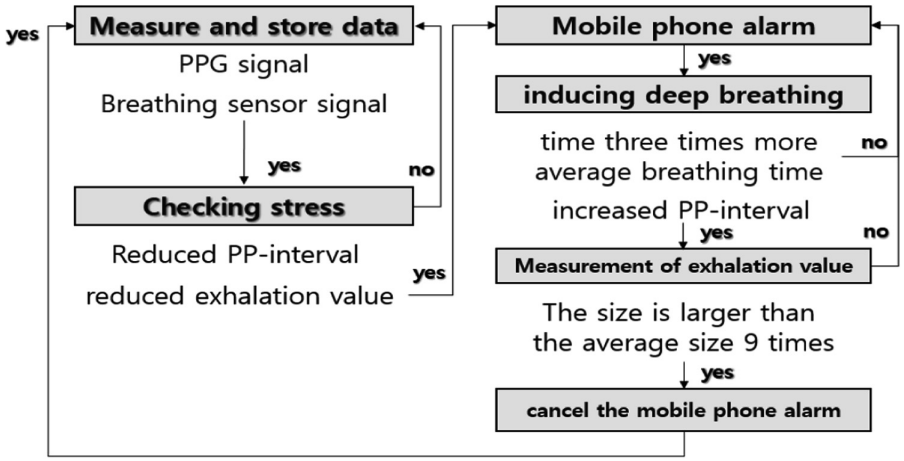


Fig. 3. Mobile health monitoring system including biofeedback training

4 Conclusion

The purpose of this study is to suggest a substantial health promotion system. Although biofeedback, which is a health promotion training method, has to be performed at the training place, such a disadvantage can be solved as a ubiquitous system is developed. The proposed monitoring system is to detect stress and breathing rate and present effective solutions. As a result of the study, the PP-interval increased during deep breathing. This indirectly shows the activity of the parasympathetic nervous system, indicating a change to a relaxed state in deep breathing, the activation of parasympathetic nerves is expected to increase the electrical stability of the myocardium and help balance the autonomic nervous system [9]. Also, the value of the expiration was increased in the breathing pattern after deep breathing. The increase in the expiration value is thought to allow more fresh Oxygen by unconsciously spewing more carbon dioxide. As a result, it increases the oxygen saturation of the blood [10]. From the two results, it was confirmed that the stress and the lack of oxygen in the cells, which inducing the generation of free radical oxygen, can be prevented.

A similar result shows that personal digital pedometer with feedback system informs actual users using smartphone showed significant effect within 24 months [11]. The cohort study has not yet been conducted in this paper. So, as a further study of this paper, we try to show the actual effect at intervals of 1 month and 2 months through the system developed according to the monitoring system. After that, we will use artificial intelligence, database technology for personalization. This is expected to improve measurement accuracy and efficiency.

The ubiquitous system using the monitoring system proposed in this study is expected to be used as a new method for improving the personal health of patients with limited movement.

Acknowledgement. This work was supported by the Dongguk University Research Fund of 2018 (S-2018-G0001-00062).

References

1. Gelogo, Y.E., Kim, H.K.: Integration of wearable monitoring device and android SmartPhone apps for u-Healthcare monitoring system. Int. J. Softw. Eng. Appl. **9**(4), 195–202 (2015)
2. Park, H., Hong, K., Kim, S.H., Shin, S.S.: Development of the practical garment apparatus to measure vital sign of ECG for U-Health care. J. Korean Soc. Clothing Text. **31**(2), 292–299 (2007)
3. Schwartz, G.E.: Biofeedback, self-regulation, and the patterning of physiological processes: by training subjects to control voluntarily combinations of visceral, neural, and motor responses, it is possible to assess linkages between physiological responses and their relationship to human consciousness. Am. Sci. **63**(3), 314–324 (1975)
4. Cadenas, E., Davies, K.J.A.: Mitochondrial free radical generation, oxidative stress, and aging. Free Radical Biol. Med. **29**(3–4), 222–230 (2000)
5. Chan, P.H., Schmidley, J.W., Fishman, R.A., Longar, S.M.: Brain injury, edema, and vascular permeability changes induced by oxygen-derived free radicals. Neurology **34**(3), 315 (1984)
6. Pickering, A.M., Vojtovich, L., Tower, J., Davies, K.J.A.: Oxidative stress adaptation with acute, chronic, and repeated stress. Free Radical Biol. Med. **55**, 109–118 (2013)
7. Akhter, N., Tharewal, S., Gite, H.: Microcontroller based RR-Interval measurement using PPG signals for heart rate variability based biometric application. In: 2015 world Congress on Information Technology and Computer Application (2015)
8. Kamath, M.V., Fallen, E.L.: Power spectral analysis of heart rate variability. A noninvasive signature of cardiac autonomic function. Crit. Rev. Biomed. Eng. **21**(3), 245–311 (1993)
9. Hee-Hyuk, L.: Effect of sports massage on post-exercise parasympathetic reactivation. Korea J. Sports Sci. **19**(1), 751–761 (2010)
10. Soon-Bok, C., Hee-Sook, K., Yun-Hee, K., Choon-Hee, B., Sung-Eun, A.: Effects of abdominal breathing on anxiety, blood pressure, peripheral skin temperature and saturation oxygen of pregnant women in preterm labor. Korean J. Women Health Nurs. **15**(1), 32–42 (2009)
11. Kelli, H.M., Witbrodt, B., Shah, A.: The future of mobile health applications and devices in cardiovascular health. Eur. Med. J. Innov. **2017**, 92–97 (2017)

A Study on Integrated Maneuvering Performance Simulation System for a Vessel

NamHyun Yoo(✉)

Department of Computer Science and Engineering,
KyungNam University, 7 Kyungnamdaehak-ro, Masanhappo-gu,
Changwon-si, GyeongNam 51767, Republic of Korea
hyun43@kyungnam.ac.kr

Abstract. Unlike cars, ship owners are able to select the model, engine type and additional options of their ship. Due to such characteristics, because ships are relatively greater in size and take longer to build, in most cases, only a few functions are verified through actual real ship test and the basic performance and maneuvering performance of a ship are verified through model ship test, analysis and simulation conducted. Moreover, since the results are provided to the ship owner in the form of a maneuvering performance booklet upon delivery of that ship, this maneuvering performance booklet is associated with all the departments in charge of design and engineering. Thus, it takes to a lot of time to build a booklet because all related maneuvering results have to be gathered from all departments and arranged consistently. This paper proposed an integrated maneuvering performance simulation system to reduce cost to gather all data from departments and to build a booklet consistently. Also, this integrated system can reinforce the research capacity of that shipbuilding company through mutually analyzing and sharing its information among each separated department.

Keywords: Maneuvering · Simulation · Modeling · Integrated management system

1 Introduction

The basic appearance and power unit of an automobile is already determined by the manufacturer and only the additional options are left to be selected by users, whereas ship owners are able to select the model, engine type and additional options of their ship. Due to such characteristics, the basic performance and maneuvering performance of an automobile are tested and improved through actual test drive. However, because ships are relatively greater in size and take longer to build, in most cases, only a few functions are verified through actual ship test and the basic performance and maneuvering performance of a ship are verified through model ship test, interpretation and simulation conducted in accordance with the international standards [1–3]. Moreover, since the results are provided to the ship owner in the form of a maneuvering performance booklet upon delivery of that ship, this maneuvering performance booklet is associated with all the departments in charge of design and engineering. In the case

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 234–239, 2020.
https://doi.org/10.1007/978-981-13-9341-9_40

where it is possible to conduct integrated management of all the design and engineering technologies, it is possible to reinforce the research capacity of the that shipbuilding company through mutually analyzing and sharing its information, and this allows that shipbuilding company to enhance its ship quality competitiveness and ship production competitiveness.

This paper proposed an integrated ship maneuvering performance simulation system capable of using the entered basic ship data and environmental data to predict the basic performance and maneuvering performance of a ship is designed and built. By using this system, it is possible to not only shorten the time it takes to design and build a ship, but also conduct integrated management of the design and engineering technologies separately owned by the different departments.

2 Design and Implementation of Integrated Maneuvering Performance Simulation System

To predict the basic performance and maneuvering performance of a ship, it is necessary to enter the ship data. In here, the ship data includes information such as length, width, height and draft. Then, it is necessary to enter the basic data related to the propeller, rudder, bridge and wave. The reason for this is because it is possible to predict the ship maneuvering performance of a ship only when the ship model is combined with the environmental information having an influence on the ship maneuvering performance. To conduct a simulation to verify the basic performance and maneuvering performance of a ship, the basic mathematical models used for calculating the motion, resistance, thrust, torque, rudder force, coriolis/centrifugal force, hydrodynamic force and wind force are required. Some of the mathematical models used in this paper are partially defined in Table 1, and the remaining mathematical models used in this thesis are defined in References. In particular, the MMG and Abkowitz of the hydrodynamic force can be suggested in [4] and [5], respectively.

Based on the entered basic data, there are 16 different types of test that must be conducted to model and simulate the basic performance and maneuvering performance of a ship. “Initial Turning Test” is a test conducted to examine the initial turning performance of a ship. While the ship is heading forward at a constant velocity, after changing the rudder angle to 10°, various data acquired until the heading angle is changed to 10°. “Course Change Test” is a test conducted to examine the turning performance of a ship. After changing the rudder angle to 15°, and when the heading angle reaches the predetermined angle (Rudder-Reverse Heading Angle), the rudder is steered to −15°, and when the heading angular velocity reaches 0, the rudder angle is changed to 0°. If the heading angular velocity reaches 0, this means that the ship’s sea route change is complete. “Turning Test” is a test conducted to examine the turning performance of a ship. By using a constant rudder angle to turn the ship in the port-starboard direction, the ship’s trajectory is acquired. According to the maneuvering performance standards set by IMO, Ad must be greater than 4.5L (Ad < 4.5L) and TD

Table 1. Mathematical models for an Integrated Ship Maneuvering Performance Simulation System

	Mathematical model	Definition
Motion	Drift Angle	If V $>$ $\in, \beta = sin^{-1}\left(\frac{-v}{V}\right)$else, $\beta = 0$
	Driving Trajectory	$d_i = d_{i-1} + \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}$
	Velocity	$V = \sqrt{u^2 + v^2}$
	Fluid Velocity (Propeller Position)	$V_\alpha = \mu(1 - \omega)$
Thrust	Quadrant 1	Quadrant 1, $T = (1 - t)\,\rho n^2 D_p^4\left(a_0 + a_1 J_A + a_2 J_A^2\right)$
	Quadrant 2	$Quadrant\ 2, T = \rho n^2 D_p^2 C_1, C_1 = a_0$
	Quadrant 3	$Quadrant\ 3, T = -\rho n^2 D_p^2 C_1, C_1 = a_0$
	Quadrant 4	Quadrant 4, $T = \frac{C_3}{C_1}(1 - t)\rho n^2 D_p^4\left(a_0 + a_1 J_A + a_2 J_A^2\right)$
Torque	Torque	$Q = \rho n^2 D^5 K_Q(V_a, n), V_a = V(1 - \omega)$
	Advance Ratio during Wind milling	$J_a = \frac{-b_1 - \sqrt{b_1^2 - 4b_0 b_2}}{2b_2}$
	Mass Moment of Inertial of Propeller Shaft Direction	$I_p = 0.006647 D_p^5 \frac{A_E}{A_o} + 0.0703 \times 1.025 \times \left(\frac{D_p^5}{\pi Z}\right)\left(\frac{P}{D}\right)^2 \frac{A_E}{A_o}$

must be greater than 5.0L (Td < 5.0L). "Accelerating Turn Test" is a test conducted to examine the turning performance of a ship during acceleration. After the ship's velocity reaches 0, the rpm is changed to 'full sea ahead' to examine the ship's turning performance during acceleration. "10/10 Zig-Zag Test" is a test conducted to examine the yaw checking ability of a ship. After changing the rudder angle to 10°, and when the course angle is changed to 10°, the rudder angle is changed to −10°, and when the course angle is changed to −10°, the rudder angle is changed to 10° again. Through measuring the primary and secondary overshoot course angles, it can be determined that the smaller the overshoot course angles are, the better the yaw checking performance is. "Pull-Out Test" is a test conducted to examine the course stability of a ship. After the rudder angle is set to 0 while the ship is turning at a constant angular velocity, the heading angular velocity is reviewed as time progresses, and after the rudder angle is set to 0 while the ship is turning in the port-starboard direction, the heading angular velocity is reviewed when the rudder angle reaches its normal status. Through comparing the heading angular velocities, it can be determined that the greater the difference is, the more unstable the course stability is. "Man Overboard Test" is a test conducted to rescue any person dropped from the boat into the sea. The main purpose is to return the boat heading forward to a certain point the boat passed within the shortest time. While the boat is heading forward, the rudder angle is changed to the maximum rudder angle in the opposite direction, and when a certain rudder-angle-changed heading angle is reached, the rudder angle is changed to the maximum rudder in the opposite direction, and when the ship turns in the opposite direction and the ship's course reaches a rudder-angle-changed heading angle close to 180°, the rudder is controlled so that the ship heads forward and returns to the point where it initially took

off. “Parallel Course Test” is a test conducted to examine the straight running stability of a ship. After the rudder angle is changed to the maximum rudder angle, and when a certain heading angle is reached, the rudder angle is changed to the maximum rudder angle in the opposite direction, and when the original heading angle is reached, the rudder is steered so that the straight running status is maintained. “Crash Astern Test” is a test conducted to examine the emergency stop performance of a ship. “Engine Stop Test” is a test conducted to examine the stop performance of a ship. While the ship is operated at the initial rpm, and when the engine stop command is given, the fuel supply to the propeller is cut off to put the propeller in the idle state and as well as to stop the ship. “Deceleration Test” is a test conducted to examine the deceleration performance of a ship. The four items to be tested are as follows: Full Sea Ahead → Full Ahead, Full Ahead → Half Ahead, Half Ahead → Slow Ahead, Slow Ahead → Dead Slow Ahead. “Acceleration Test” is a test conducted to examine the acceleration performance of a ship. While the ship is stopped, the rpm is changed to ‘Full Sea Ahead’ to measure the time and track reach it takes for the rpm to reach ‘Full Sea Ahead’. “Turning Circle Test in Shallow Water” is a test conducted to examine the turning performance of a ship in shallow water. The ship is turned at a constant rudder angle for a set amount of time in the port/-starboard direction to acquire the ship’s trajectory. “Squat Test” is a test conducted to examine the trim variation and draft variation of a ship in shallow water. The ship’s velocity is increased under various HD conditions to measure the trim angle variation and draft increase. “Course Keeping Limitation Test” is a test conducted to find out the rudder angle and drift angle required for a ship to maintain its normal status when influenced by wind load. Lastly, “Minimum Ship’s Speed Making Headway” is a test conducted to find out the minimum velocity required for a ship to head forward. While the fuel supply to the engine is cut off and the ship is still heading forward, the velocity required for maintaining the sea route is measured. The results generated through simulation test consists of records saved as a text file, and the recorded information includes the ship’s trajectory based on the ship’s maneuvering motion as well as the ship’s location and time based on each condition. The fundamentally saved information includes X trajectory, Y trajectory, course angle, propeller rev count, track reach, yawing angular velocity, drift angle and rudder angle. Although each module that combines such information to display the maneuvering performance on a real-time basis had its own functions applicable to each type of test, as for the ship’s trajectory and velocity variation, a common module was implemented and utilized.

Using N_Main as the standard, the integrated ship maneuvering performance simulation system can be divided into DBAdapter, NASIMUL, NAReport, Input Basic Data, User Management, DB Management, Booklet Data Management, and Printer Management. DB Adapter and DB Management are modules for database and interface. In particular, although this system was normally operated based on Oracle, when a test was conducted on the sea, the MS Access database was used to prevent network disconnections. To ensure data integrity, the adapters were configured in a way mutually compatible and were processed to cause no errors during data conversion. User Management managed the user schema section to allow all the personnel in charge of design and engineering to mutually exchange and utilize their information. In addition, this component was connected to DB Adapter and Management to configure

two different modes: simultaneous user mode and single user mode. Input Basic Data was the section that received the following information: ship data, environmental information and other information. NASIMU, the core section of engineering, played the role of directly conducting 16 different types of test. NAReport was a tool used to visualize the maneuvering performance of a ship based on the text data calculated by NASIMUL. Lastly, Booklet Data Management played the role of using the data generated from NASIMUL and NAReport to generate a maneuvering performance booklet for particular ships with one button press through the mail merging function of MS-Word. Figure 1 shows a class diagram of Integrated Maneuvering Performance Simulation System.

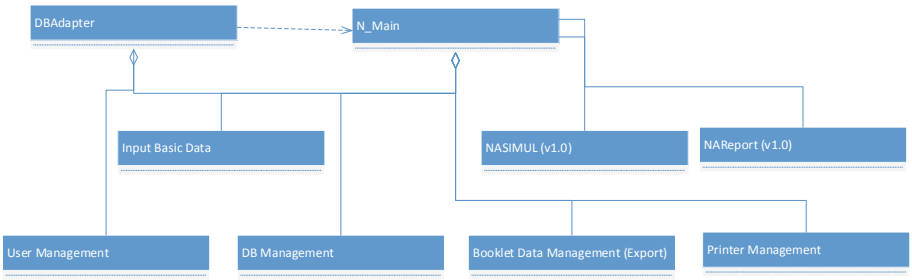


Fig. 1. A class diagram of Integrate Maneuvering Performance Simulation System

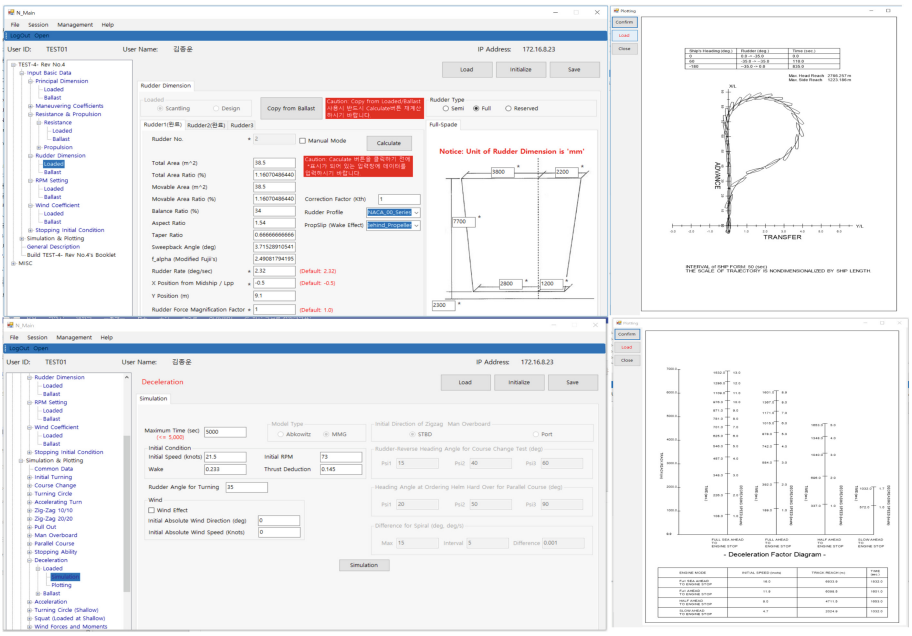


Fig. 2. Screenshots of Integrate Maneuvering Performance Simulation System

C# and C++ were the two development languages simultaneously used to develop the integrated ship maneuvering performance simulation system, and the modules developed through these two languages exchanged data through marshalling.

The reason for running such development process is as follows: because the engineering core was developed by an internal developer who worked as an engineer at the shipbuilding company, and because the user interface section and visualization section were developed by an external developer, it was necessary to prevent reverse engineering from exposing the core algorithm required for conducting a ship maneuvering performance simulation in the simultaneous development process. Figure 2 shows a screen shot of Integrated Maneuvering Performance Simulation system implemented in this paper.

3 Conclusion

As described earlier, because ships are relatively greater in size compared to automobiles and are decided mainly by ship owners, to shorten the time it takes to design and build a ship, only a few functions of a built ship were evaluated by real ship test and the remaining functions were either verified through model ship test, analysis and simulation in accordance with the IMO standards or verified through utilizing a simulation system constructed based on the experience acquired through years of real ship tests conducted on the sea. However, since each information was individually managed by each department, the management process itself was neither systematic nor efficient. The development and application of this system made it possible for each department to share its individually owned technology and information with other departments, and this prevented researchers from conducting overlapping tests. In addition, the development and application of this system made it possible to share other users' tests, and this significantly shortened the overall lead-off time required in the design phase and engineering phase.

References

1. Pakkan, S.: Modeling and Simulation of a Maneuvering Ship. Middle East Technical University (2007)
2. Abdel-latif, S., Celiel, M., Eldin Zakzouk, E.: Simulation of ship maneuvering behavior based on modular mathematical model. In: 21st Mediterranean Conference on Control and Automation, Crete, pp. 25–28 (2013)
3. Stern, F., Agdrup, K., Kim, S.Y., Hochbaum, A.C., Quadvlieg, F., Perdon, P., Hino, T., Broglia, R., Gorski, J.: Experience from SIMMAN 2008 – the first workshop on verification and validation of ship maneuvering simulation methods. J. Ship Res. **55**(2), 135–147 (2011)
4. Abkowitz, M.A.: Measurement of Hydrodynamic Characteristics from Ship Maneuvering Trials by System Identification, New Jersey (1980)
5. Yasukawa, H., Yoshimura, Y.: Introduction of MMG standard method for ship maneuvering predictions. J. Mar. Sci. Technol. **20**(1), 37–52 (2015)

A Study on Debugging Method for Engineering Software Using Marshalling in Restricted Environment

NamHyun Yoo(✉)

Department of Computer Science and Engineering,
KyungNam University, 7 Kyungnamdaehak-ro, Masanhappo-gu, Changwon-si,
GyeongNam 51767, Republic of Korea
hyun43@kyungnam.ac.kr

Abstract. Although such commercial products such as ANSYS, NASTRAN and PATRAN provide most-wanted and productive functions used for analysis and design, they do not provide particular functions required by a manufacturing company. In this case, most companies develop the required engineering software on their own, and use C++ or C Language to develop the required engineering software. Recently, as the preference for GUI increases, the demand for GUI development increases as well. However, it is difficult for CUI developers to adapt to GUI development in a short period of time. In this paper, a debugging method for cooperation between internal developers who develop engineering software and external developers who develop GUI was proposed. This debugging method is advantageous in that it makes possible to develop GUI-based software without making significant changes to the basic structure of CUI-based engineering software. In addition, this debugging method is advantageous in that the core technology of a company can be protected, since external GUI developers cannot know the internal properties of engineering software.

Keywords: Engineering software · Cooperative debugging · Marshalling · Technology protection

1 Introduction

Most companies make efforts to develop products that demonstrate price/quality competitiveness in the market by investing research costs and personnel. To do so, design and engineering are considered essential. In particular, most frequently utilized engineering software for design and engineering in the machinery industry are ANSYS [1], NASTRAN [2], and PATRAN [3]. Engineers can utilize the engineering software to conduct diverse experiments and interpretations. Most of the engineering software are commercial products. Although they provide diverse and most-wanted functions, they rarely provide functions specially required by each company. In the case where pre-existing commercial software do not provide functions preferred by a company, that company orders its engineers to develop in-house engineering software. In this case, the development is executed through traditional engineering-oriented advanced

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 240–246, 2020.
https://doi.org/10.1007/978-981-13-9341-9_41

languages such as CUI-based C, C++ and FORTRAN. This is because engineers with engineering knowledge are mostly able to interpret the results made of nothing but texts. Recently, as the preference for GUI increases, the efforts are made to develop GUI. However, engineers require additional efforts and time to be able to develop GUI. The problem with outsourcing GUI development to an outside company is that the core engineering technology owned by a company can be exposed.

In this paper, an effective development method that reinforces security was developed for cooperation between internal developers who conduct engineering and external developers who conduct GUI development for the functional expansion of engineering software. This method is advantageous in that it allows companies to protect the core technology of their individually owned CUI-based engineering software while efficiently improving the user-friendly functions of their software.

2 Related Work

2.1 Development Methodology

Most people are not aware that the process of developing software is identical to the process of designing and constructing a building. They misunderstand such process as an easy job that can be performed by anyone. However, the process of developing software is just as complicated as the process of constructing a building, and there are diverse development methodologies. Waterfall Model [4] is a traditional software development methodology and Agile is a development methodology which improved the problems demonstrated in Waterfall Model. Because traditional development methodologies require many procedures and a considerable number of documents to support such procedures, Agile is a coding-oriented development methodology proposed to improve the unpredictable and difficult-to-manage problems demonstrated in unplanned development methodologies [5, 6]. In this paper, a development methodology to which a modified version of Agile is applied was utilized.

2.2 Marshalling

Marshalling is a method used to implement RPC (Remote Procedure Call) required for data transfer between processes or threads, and is used for conversion between non-controlled type and CLR type under the P/Invoke process in Microsoft's Com and .Net framework. In this paper, this technique was used to not only wrap internal developers' engineering algorithm by converting it into a DLL file, but also transfer data with external developers' user interface program. In the case where internal developers develop development source codes, debugging becomes possible through the setting of a development tool. However, this paper was limited to the case where a DLL file is provided by reinforcing the security of the company that internal developers work for.

3 Improvement of an Existing Engineering Software

3.1 An N Software

An existing engineering Software, called N software using the debugging development method proposed in this paper predict the maneuvering performance of a ship when the following information are entered: ship data and environmental information such as tidal current flow and wind speed. The N software is an integrated management system that integrates and manages all results in one application program instead of simply listing the engineering algorithm results owned by each department in one report.

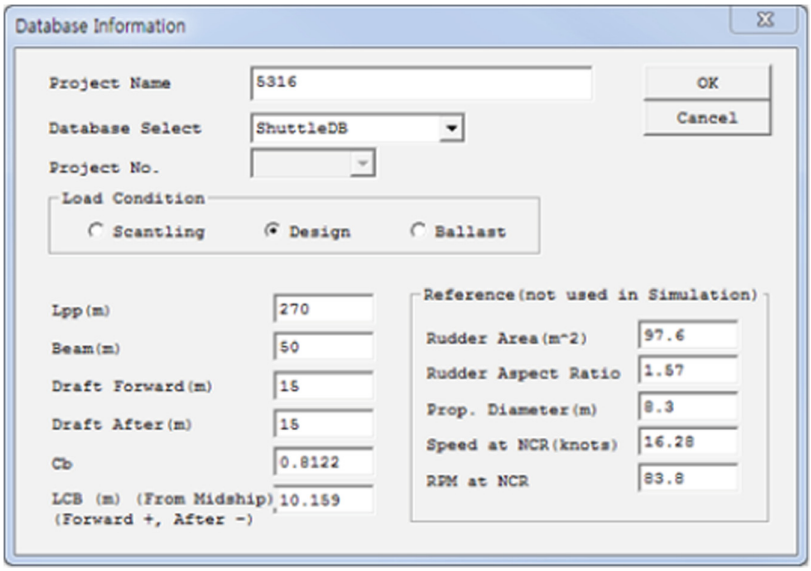


Fig. 1. A screen shot of N Software owned by D company

N software is an in-house engineering software developed by Company D in the early 2000's, and the main purpose of this engineering software is to provide the complete maneuvering performance of a ship to the ship owner who orders that ship. Although N software contains the basic GUI, it requires users to enter and save the database information by individually loading its 6 input boxes, and its 16 interpretation simulation processes require users to load a separate data input box, enter and save data, run the simulation, and repeat this process until preferred results are acquired. Users are required to save the final simulation results as a separate image file and insert this image file into a specific report. The problems with this engineering software is that it normally takes 2–4 h to complete one result report. Figure 1 is a screen shot of the N software owned by Company D.

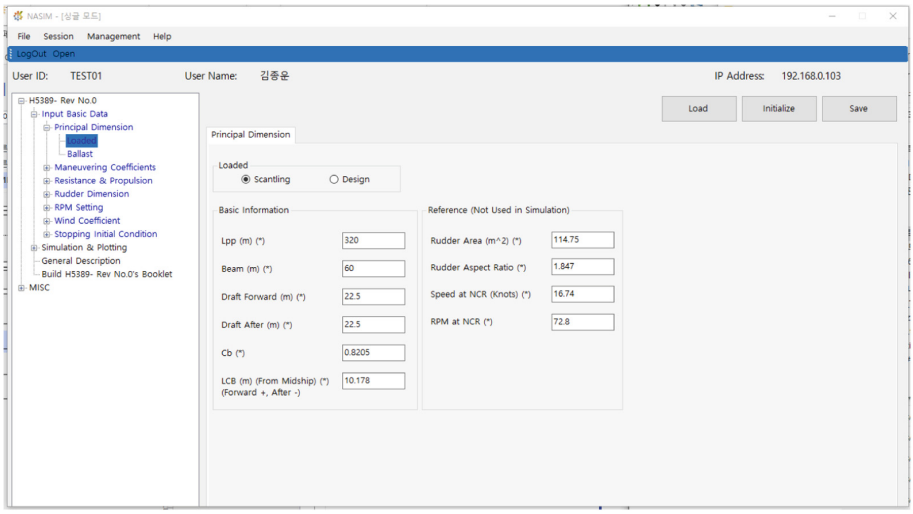


Fig. 2. A screen shot of improved user interface of the N Software based on separation between engineering logic and user interface.

3.2 Separation Between Engineering Logic and User Interface

Since N Software was developed by an internal developer who did not specialize in software engineering, it was based on simple CUI-based algorithms. And since its program logic and user interface were not separated, it was difficult to conduct maintenance and updates, and the initial developer was the only person capable of conducting debugging and performance improvements. In this paper, to improve the weaknesses of this engineering software, the engineering algorithm and user interface of this engineering software were separated through the internal developer who worked at Company D. In addition, a tree-structured interface was applied to the simple list-type data input section and simulation section in order to make it possible to directly confirm the current data input status and simulation progress status. In addition, conveniences were added to this engineering software through the mail-merge function which allows users to use one command to complete a result report. It takes to complete a final report was shortened to 1/4–1/5. Figure 2 show a screen shot of the user interface improved through this thesis. To allow users to confirm the input status on a real-time basis, it was modified so that the menu names listed in the left tree become blue when data are entered and become black when data are not entered.

3.3 Application of Marshalling

The internal developer from Company D prevented the core engineering logic from being exposed to the outside through reverse engineering, converted it into a DLL file for data transfer with the user interface program developed by an external developer, and determined the specific details of the data structs for information exchange by reaching an agreement with the external developer. In the user interface program,

"S_ACCEPT_DATA", a struct for transfer of engineering logic data, and "S_ACCEPT_DATA", a struct for transfer of result data were defined, and the marshalling technique was used for mutual data transfer between the user interface program developed with C# and the DLL file developed with C++.

4 Application of Debugging Method in Restricted Development Environment

4.1 Restricted Development Environment and Problem

The internal developer from Company D and the external developer developing the user interface had a few limitations in the process of developing the program. The main problem was that, due to the security policy of the company that the internal developer works for, the information exchange between the internal developer and external developer was limited. In addition, since the internal and external developers were unable to participate in the development process within the same physical space, some of the occurring problems were as follows.

Initially, in the process of modifying source codes, when the internal and external developers exchanged data after accidentally forgetting to modify any part of the data. An error message explaining that the memory domain is invaded appeared, and the program was shut down. In this case, because the developers were unable to confirm which part of the mutually shared memory domain was invaded, they had to go through the process of separately checking and then mutually comparing their development codes.

Secondly, in the process of transferring input data to the engineering program through the user interface program, there was a case where any part of the data was incorrectly converted and then transferred. For example, in the process of selecting a check box, '0' had to be transferred when '1' was not selected. However, when such data value was transferred backwards, whether the simulation results generated based on that data value were correct simulation results transferred based on correct simulation results or were incorrect simulation results transferred based on incorrect simulation results was not confirmed.

Thirdly, as a problem within the engineering problem, when the engineering algorithm itself was incorrect and incorrect result values were generated by the incorrect engineering algorithm, because it was impossible to trace the value variation of the involved variables, it was difficult to confirm whether or not the involved algorithm had problems. In particular, in this case, because the internal and external developers experienced difficulties in finding each other's mistakes, they also experienced difficulties in the debugging process.

4.2 Applying of Debugging Method for Restricted Development Environment

The debugging method designed and applied in this paper was named as "Marshalling Debugger", and its architecture is as shown in Fig. 3.

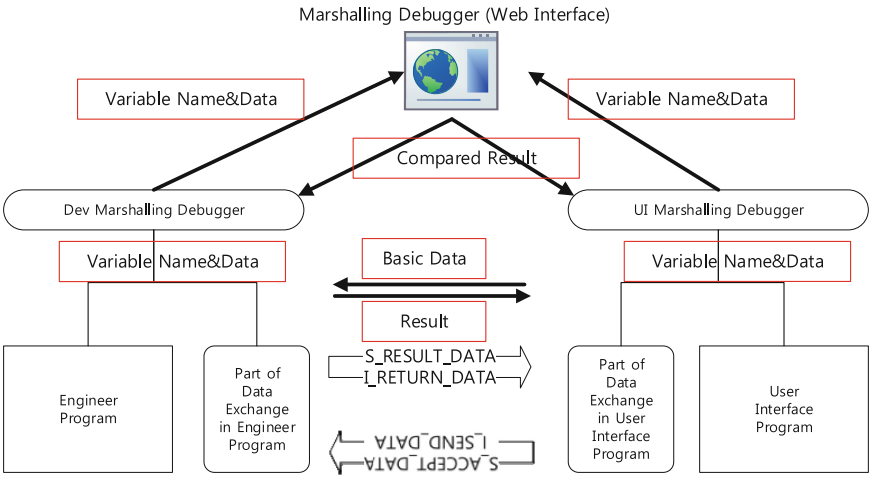
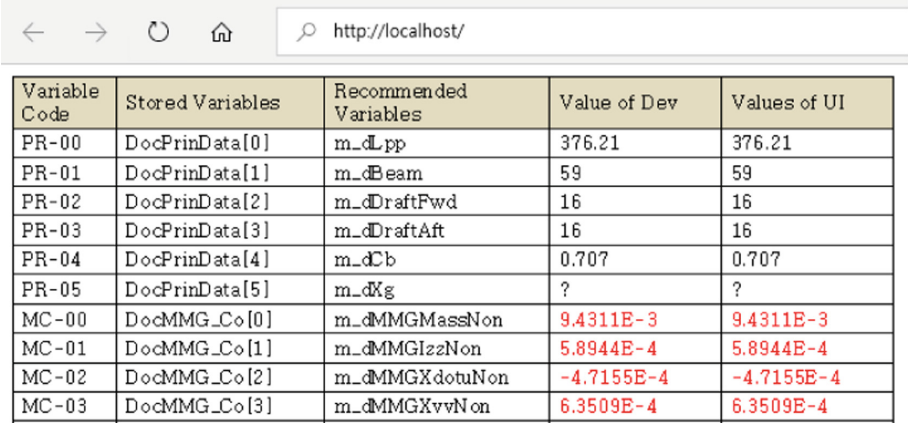


Fig. 3. An architecture of a Marshalling Debugger



Variable Code	Stored Variables	Recommended Variables	Value of Dev	Values of UI
PR-00	DocPrinData[0]	m_dLpp	376.21	376.21
PR-01	DocPrinData[1]	m_dBeam	59	59
PR-02	DocPrinData[2]	m_dDraftFwd	16	16
PR-03	DocPrinData[3]	m_dDraftAft	16	16
PR-04	DocPrinData[4]	m_dCb	0.707	0.707
PR-05	DocPrinData[5]	m_dXg	?	?
MC-00	DocMMG_Co[0]	m_dMMGMassNon	9.4311E-3	9.4311E-3
MC-01	DocMMG_Co[1]	m_dMMGIzzNon	5.8944E-4	5.8944E-4
MC-02	DocMMG_Co[2]	m_dMMGXdotuNon	-4.7155E-4	-4.7155E-4
MC-03	DocMMG_Co[3]	m_dMMGXvvNon	6.3509E-4	6.3509E-4

Fig. 4. A result of a Marshalling Debugger

The architecture of Marshalling Debugger consists of "Dev Marshalling Debugger" connected to the engineering program, "UI Marshalling Debugger" connected to the user interface program, and "Marshalling Debugger" tracing the data value of variables. The data exchange among "Dev Marshalling Debugger", "UI Marshalling Debugger" and "Marshalling Debugger" is executed based on Web interface, and the scope of traced variables is determined according to the data set within the user interface program. In general, only the changes made to the data within the mutually exchanged data section are monitored and transferred to "Marshalling Debugger". However, when users want to trace the variables within the program, each specific method can be

assigned with a serial number, and the details of the variables to be traced within the method within that serial number can be displayed. Figure 4 shows a result of "Marshalling Debugger" through web interface.

5 Conclusion

The engineering software development environment of Company D that executed a cooperative development process through this paper is an environment that can be frequently seen in the machinery industry. In this case, most companies order pre-existing developers to be able to develop GUI or provide their development source codes to an outside company as long as it is not a problem to have the involved software exposed. However, in the case where the involved algorithm cannot be exposed to competitors or external developers, since it is not possible to expose the involved algorithm to external developers, most companies wait until their internal developers become able to develop GUI or continue to use their pre-existing program accompanied by inconveniences. "Marshalling Debugger" proposed in this paper is advantageous in that it can be used to exchange between input data and engineering result data without having to expose the internal algorithms that need to be secured, and in that internal developers can cooperate with external developers to run the development process without having to learn how to develop a program capable of displaying result data in 2D and 3D.

References

1. ANSYS. http://www.ansys.com/
2. NASTRAN. http://www.mscsoftware.com/product/msc-nastran/
3. PATRAN. http://www.mscsoftware.com/product/msc-patran/
4. Royce, W.W.: Managing the development of large software systems: concepts and techniques. In: 9th International Conference on Software Engineering. IEEE Press, Monterey (1987)
5. Abrahamsson, P., Salo, O., Ronkainen, J., Warsta, J.: Agile Software Development Methods: Review and Analysis. VTT Publication, Finland (2002)
6. Cao, L., Ramesh, B.: Agile Requirements engineering practices: an empirical study. IEEE Softw. **1**, 60–67 (2008)

Directional Antenna Prediction Control System for Stable Communication of Small Unmanned Surface Vehicle

NamHyun Yoo(✉)

Department of Computer Science and Engineering,
KyungNam University, 7 Kyungnamdaehak-ro, Masanhappo-gu, Changwon-si,
GyeongNam 51767, Republic of Korea
hyun43@kyungnam.ac.kr

Abstract. Although Wi-Fi has the best versatility and provides relatively fast bandwidth, it has problems that the obstacle transmittance is low and the communication speed above a certain level is ensured only when LOS is guaranteed. MR-SENTINEL, a marine unmanned ship transmits image information for remote control of an unmanned ship and monitoring of maritime environment by using WiFi directional antenna in order to reduce operating cost and to extend coverage. However, since a small marine unmanned ship has frequent Pitch, Roll, Sway situation due to wave, when a directional antenna is used, it is necessary to cope with the pitching and rolling through a horizontal balancing device such as a jimbol to provide a stable bandwidth, and an antenna rotator is required to maintain the direction of the directional antenna during the situation of Sway and the movement. In this paper, we designed and implemented a directional antenna predication control system to provide a stable communication environment of WiFi directional antenna used in small unmanned ships and verified the performance through field experiments.

Keywords: Unmanned ship · WiFi · Directional antenna · 6 degrees of freedom · Balancing system

1 Introduction

As the drone industry has been activated in recent years, various kinds of drones have been released and applied in various fields. The most active fields are aviation and land drones, which are already used with Predator in the field for the military purpose. In addition, to apply them to the shipping field, research and policy establishment has been conducted in order to introduce the space-based lot number system from the plane-based lot number system.

Recently, private companies such as DJI have been selling personal drones in addition to enterprises. On the other hand, the marine sector is relatively slower than the aviation and land sectors, and is mostly limited to military, public, and research fields. It is true that there are limitations due to law and accessibility, but this is due to the fact that it is necessary to use relatively expensive satellite communication or 4/5th generation communication in order to operate the marine drones. In order to solve these

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 247–252, 2020.
https://doi.org/10.1007/978-981-13-9341-9_42

problems, Wi-Fi, which supports a communication distance of more than 1 km in [1] or RF (Radio Frequency) used in RC (Radio Control) is often used. RF can give simple control commands or transmit only low quality images. Although Wi-Fi has the best versatility and provides relatively fast bandwidth, it has problems that the obstacle transmittance is low and the communication speed above a certain level is ensured only when LOS is guaranteed. MR-SENTINEL (Marine Robot-Sentinel), a small marine unmanned ship proposed in [2] transmits image information for remote control of an unmanned ship and monitoring of maritime environment by using WiFi directional antenna in order to reduce operating cost.

However, since a marine unmanned ship has frequently faced Pitch, Roll, Sway situation due to wave, when a directional antenna is used, it is necessary to cope with the pitching and rolling status through a horizontal balancing device such as a jimbol to provide a stable communication, and an antenna rotator is required to maintain the direction of the directional antenna during the situation of Sway and the movement. In this paper, we designed and implemented a directional antenna predication control system to provide a stable communication environment of WiFi directional antenna used in small unmanned ships and verified the performance through field experiments.

2 Related Work

Basically, communication systems for unmanned ships such as drone use Bluetooth, WiFi, cellular communication, satellite communication, and LTE (Long-Term Evolution). Recently, research is being conducted to apply the 5th generation mobile communication system [3].

In DARPA (Defense Advanced Research Projects Agency), a military unit is studying and developing a program called Mobile Hotspots to use drones in the field. Using this system, it is possible to implement a PAT (Pointing, Acquisition, Tracking) function that can direct, acquire, and track the drones in the front through interconnections [4]. Google is expanding its wireless Internet supply network by using drones through a company called Titan Aerospace and testing data transmission technology using phased array at the same time. Facebook is also working on a similar project. Intel and AT&T confirmed the possibility of a 5G communication system using crowd drones at the Sydney Opera House and the opening ceremony of the 2018 Pyeongchang Olympic Games. Basically, however, they are developing a drone-related communication system using LTE and IoT (Internet of Things). China Mobile, together with Ericsson of Sweden, applied the 5G communication system to drones for the first time in the world [5].

In Korea, LTE network is mainly used to use drones. KT has developed a drone LTE communication service to provide LTE communication services by using crowd drones in the disaster area. LG U + and SK Telecom are also developing and distributing drone control and utilization services using LTE. In addition, LTE, LPWAN (Low-Power Wide Area Network), and Wi-Fi are used to research, develop and commercialize communication systems for drones [5].

3 MR-SENTINEL

MR-SENTINEL proposed in [2] is a system developed to be used by fishermen and in fishing villages that want to monitor fishery raising and marine environment or for personal hobby purpose, rather than being developed for military, research, and public purposes. Therefore, MR-SENTINEL was developed mostly using common parts such as a gyro sensor, IP cameras based on an embedded PC, not expensive parts such as IMU (Inertial Measuring Unit). Figure 1 shows an unmanned ship system applying the directional antenna prediction control system designed and implemented in this paper. The hull of MR-SENTINEL is catamaran, and two electric motors are used as propulsion for forward/backward and direction change. Steering is done with two joysticks.

MR-SENTINEL can monitor the remote control within about 1 km in an obstacle-free environment between the remote control system and the unmanned ship and can transmit a video over 1280p in real time. One of the cameras is used to detect obstacles in front during moving while the other is used to detect garbage floating on the sea surface or jellyfish, a harmful marine organism.

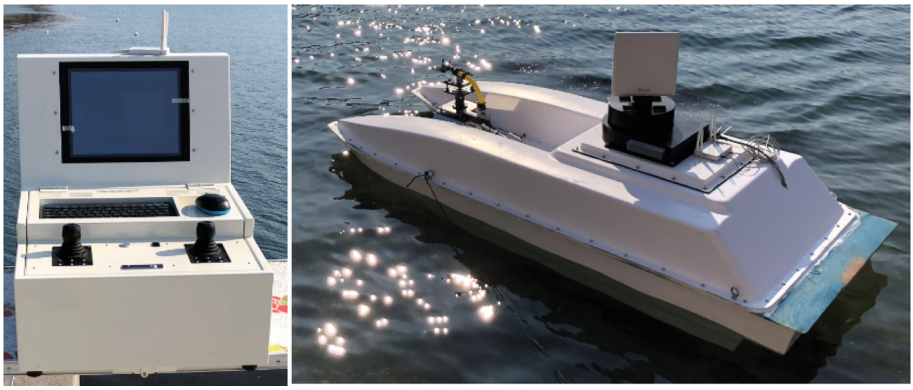


Fig. 1. Remote control system and small unmanned surface vehicle, is called MR-SENTINEL

4 Directional Antenna Prediction Control System

The directional antenna prediction control system largely provides an optimum communication environment system while the unmanned surface vehicle moves by the remote control.

The directional antenna used in this paper is F model of F-company. When the pitching or rolling situation occurs more than 10º above and below the MR-SENTINEL with this model, MR-SENTINEL enters the dead-reckoning state due to low of the hull length and height. When entering the dead-reckoning state, MR-SENTINEL will use the automatic mode to escape the situation.

In the movement or Sway situation, the directional antenna prediction control system maintains communication with the remote control system through the rotation of the antenna motor. Figure 2 shows the purpose of directional antenna predication control system for MR-SENTINEL.

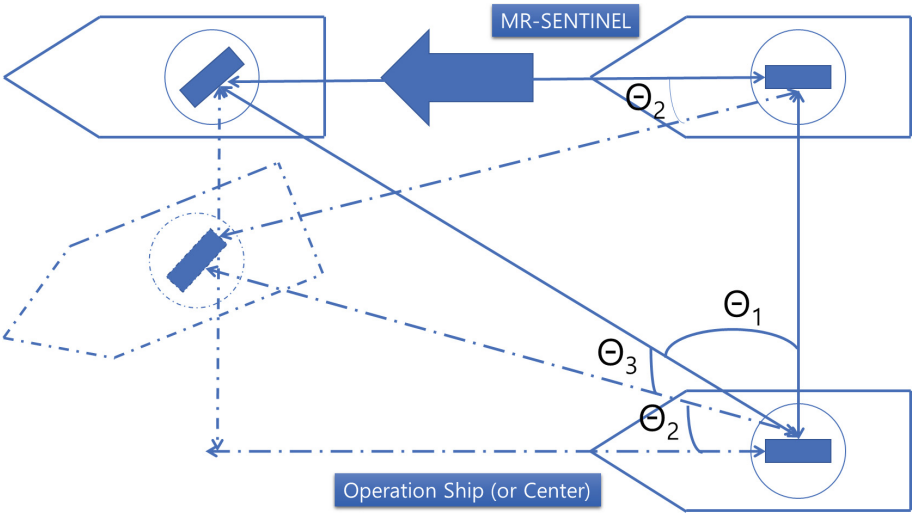


Fig. 2. The purpose of directional antenna prediction control system at sea

$$S_{scale} = \frac{S_{std}}{S_{maxrpm}}$$

$$RP_{rpm} = RP_{value} - S_{std}, RP_{rpm} = \frac{RP_{rpm}}{S_{scale}}$$

$$LP_{rpm} = LP_{value} - S_{std}, LP_{rpm} = \frac{LP_{rpm}}{S_{scale}} \quad (1)$$

$$Diff_{rpm} = RP_{rpm} - LP_{rpm}$$

$$\text{if } Diff_{rpm} > 0, \Delta\theta = \theta_{origin} - \alpha Diff_{rpm}$$

$$if\ Diff_{rpm} < 0, \Delta\theta = \theta_{origin} + \alpha Diff_{rpm}$$

In order to control the antenna, the expected direction of MR-SENTINEL should be known by the thrust of each propeller.MR-SENTINEL. Equation (1) represents the process of calculating the expected angle required for the directional antenna control motor. In this Eq. (1), S_{ratio} is the ratio of the moving distance of the joystick to the RPM of the motor, S_{maxrpm} is the maximum RPM of the motor with respect to the

maximum moving distance of the joystick, RP is the right propeller, LP is the left propeller, RP_{rpm} is the RPM of the right propeller, RP_{value} is the motion data from the right joystick, $Diff_{rpm}$ is the different RPM between the two propellers, and α is the direction coefficient. In this equation, external disturbances such as currents, winds, and waves were not considered. $\Delta\theta s$ is the moving angle of the motor to be moved in advance, and it is proportional to the moving speed of the ship.

Fig. 3. The experiment of using directional antenna prediction control system at sea

Figure 3 shows the experiment of using directional antenna prediction control system at sea. Figure 4 shows the experimental results when the directional antenna prediction control system is used or not used. In the left figure, the directional antenna control system is used, and it is confirmed that the communication connection is continuously ensured without entering the dead reckoning state. The right figure shows the case where the directional antenna prediction control system is not used and the system enters the dead reckoning state and returns. In severe cases, it may not be possible to return.

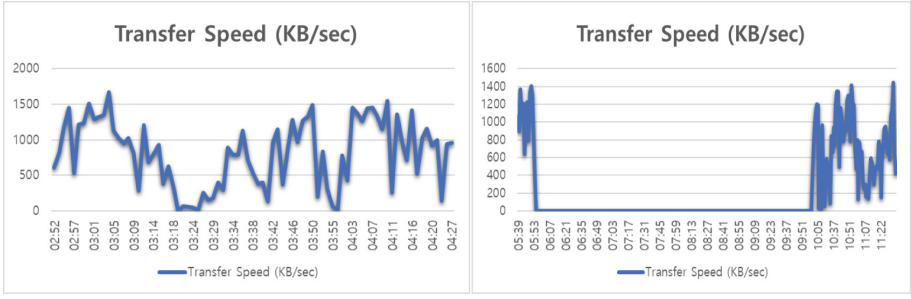


Fig. 4. Left figure shows the experimental results when the directional antenna prediction control system is used and right figure shows the results is not used.

5 Conclusion

The directional antenna prediction control system designed and implemented in this paper used a WiFi directional antenna to remotely control small unmanned surface vehicle and to transmit image information for maritime surveillance. Basically, the WiFi directional antenna provides the maximum angle for the communication in horizontal and vertical environments through the BeamWidth technology specification. The directional antenna used in this paper is relatively narrow (Horizontal 9° and vertical 11°) and often enters the dead-reckoning situation due to the influence of the waves generated by the movement of large ships around or in an influence of waves over 1 m even if the directional antenna prediction control system implemented in this paper is applied. Therefore, we decided to use a non-directional antenna instead of the existing directional antenna in MR-SENTINEL. If there are no obstacles within the maximum 1–2 km, this antenna provides about 80–90% performance compared with the conventional directional antenna described in this paper. However, the use of a directional antenna has the advantage of widening the range of areas using WiFi for signal transmission between antennas at fixed positions and intermediate signal amplification for non-directional AP.

References

1. Raman, B., Chebrolu, K.: Experiences in using WiFi for rural internet in India. IEEE Commun. Mag. **45**(1), 104–110 (2007)
2. Yoo, N.H.: ZIB structure prediction pipeline: design and development of small unmanned surface vehicle system for marine surveillance based on remote control. In: 7th International Conferences on Advances in Next Generation Computer and Information Technology, pp. 180–184. SERSC, Hokkaido (2018)
3. Kim, H.W., Kang, K.S., Chang, D.I., Ahn, J.Y.: Technical and Standardization Trends on Control and Non-Payload Communications for Unmanned Aircraft System: Electronics and Telecommunications Trends, ETRI (2015)
4. Mobile Hotspots. https://www.darpa.mil/program/mobile-hotspots
5. Jeon, H.S.: Technology Trend of Drone Communications: Weekly ICT Trends, IITP (2017)

Design of a Full Body Tracking-Based VR Fighting Game

Hoyong Kim, Guoqing Wu, and Yunsick Sung(✉)

Department of Multimedia Engineering,
Dongguk University-Seoul, Seoul 04620, South Korea
sung@mme.dongguk.edu

Abstract. With the widespread distribution of Virtual Reality (VR) devices, the number of VR users and the size of the VR contents market are steadily increasing. In response to the users' demand, full-body tracking technologies using additional peripheral devices have been developed. Games using such technologies have begun their development as well. In order to make the games more immersive, it is imperative to develop further technologies related to full-body tracking. This paper involves the design of a VR game supporting full-body tracking. Furthermore, additional systems that are implemented in the game are also developed by designing a new submotion system. The proposed submotion system recognizes the user's motion by dividing it into small sub-motions. Based on this, this paper proposes advanced damage and attacking systems. In addition to making the games more fun and immersive, the proposed system will allow for the research and development of an engine applicable to full-body tracking VR contents.

Keywords: Virtual reality · Fighting game · Submotion

1 Introduction

Due to the price reduction of VR devices and release of low-priced consumer-level VR devices, the number of VR users is steadily increasing. In response to the increased user demands, various new VR contents are being released as well. Using devices such as VIVE Trackers, even users are able to experience full-body tracking, which is capable of reading the motions of the entire body [1]. However, there is a limited amount of contents based on full-body tracking available to the users. VR motion recognition technology also needs to be further improved. These reasons illustrate the need for a new and improved VR motion recognition system, and a game that properly supports full-body tracking.

In a VR game, tracking the user's motion can be divided into two different methods. In the first method, the motions of a user wearing a tracking device are directly captured and translated into a game [2]. Since the motions are not predetermined, the user is given much freedom in movement. However, there is an issue with recognizing and processing the intention behind such movements. The second method recognizes the user's movement, and outputs a pre-programmed animation corresponding to the movement [3, 4]. Such process limits the movement. If the user's

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 253–258, 2020.
https://doi.org/10.1007/978-981-13-9341-9_43

movement is not precisely measured and processed, the user may find the VR content immersion-breaking and lose interest, making the content very short-lived. Furthermore, existing full-body tracking VR contents often interpret the movement erroneously as something that is not as quick and precise.

In this paper, a VR fighting game which supports full-body tracking using VR devices and dedicated tracking equipment is designed. A new VR motion system which is an improvement upon the existing system is also developed and implemented into the game. The system does not only rely on HMD and controller, but also reads the entire body's motion using the tracking devices. It is designed in such a way that ensures compatibility with various types of motion capture equipment. In this paper, VR technologies for developing the immersive, practical VR contents are proposed, which are also expected to serve as a foundational technology upon which various full-body tracking applications may be built.

2 VR Fighting Game

In this chapter, the proposed VR fighting game is introduced. Specifically, the development environment and the characteristics of the game are discussed. The game is a 1-on-1 fighting game. Like most other fighting games, the player can inflict damage upon the opponent with a successful hit. Once the accumulated damage goes over a certain threshold, the opponent is defeated, and the user wins the game. This game uses VR devices and full-body tracking. The user's movement is matched directly onto the in-game character.

The game is developed using Unity 3D. Unity 3D is suitable for VR game development, as it supports XR for VR and provides various plugins such as Unity Asset Store for SteamVR plugins which are highly useful for VR content development. HTC Vive is used for the HMD mount, and the VIVE Tracker is used for movement tracking. Vive Tracker is an additional peripheral device that can be used with HTC Vive. It has relatively quick feedback, and has moderate tracking accuracy. However, since the number and location of Trackers that can be used together are limited, further processing is required to enable full-body tracking.

The game uses VR input devices. VR devices allow for accurate measurement of various factors such as movement velocity and motion size, which means they can distinguish between similar movements and apply different effects on each of them. For example, simply stretching an arm can produce different punching motions depending on the arm's angle, location, and the relative position of other body parts. Such characteristics of VR are utilized in order to remedy the existing motion recognition methods' shortcomings.

3 Submotion System

The Submotion system is a new system which can be applied not only to full-body tracking VR contents, but too many different VR contents in general.

3.1 Overview of Submotion System

The proposed game reads the user's actual movement in short segments, categorizes them into submotions, and then produces an output. Based on the categorized submotions, various factors such as VR character's animation, direction, and velocity are produced differently. Aside from the motion capture and animation, this system can also be used for damage, effect, and the target's animation when damaged. In this paper, a new VR motion recognition system called the submotion system is built in the design of the proposed game.

For instance, when using the existing motion recognition methods, the movement of a user's forwarding the arm is recognized as one big punch motion and output it accordingly. However, in this system, the user's motion is divided into preparation, pulling the arm backward, and forwarding the arm, which form a single punch motion. Such motion recognition in a short unit makes it possible to lead the movement of the user's forwarding the arm not necessarily to the same punch motion but to a fake motion if the user stops in the middle or an uppercut if the user raises the arm. Thus, it is possible to choose various motions by altering the animation in the middle through various branches of such motions (Fig. 1).

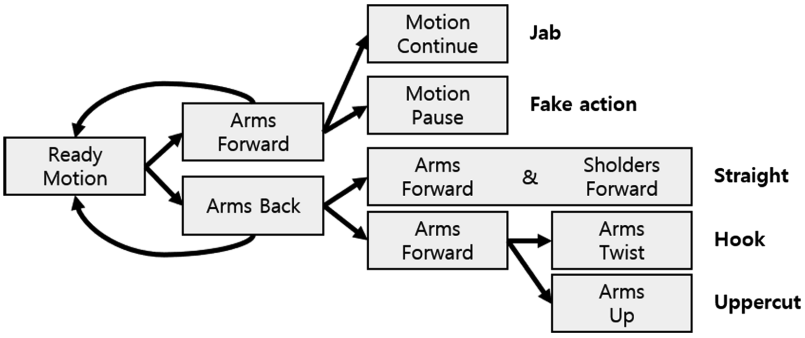


Fig. 1. Branching by submotion classification.

3.2 Design Details

In order to implement the submotion system, we conduct detailed design of three successive stages such as the user's movement measurement, classification into submotions, and classified movement processing.

First, the values measured by sensors are obtained and the change of the user's movement is measured using these measured values. We can not only measure the 3-dimensional position of the sensors through the sensors of the controller and the trackers worn by the user but also find out elements for measuring the movement such as velocity, acceleration, and movement direction of the sensors. This enables us to recognize the user's movement instead of the simple position change of the sensors according to the measured elements and the parts where the user wears the devices.

Second, the submotion most similar to the measured user's movement is found and classified. We define the submotions according to the measured user's movement in advance. The difference between the current possible next submotions and the measured user's movement is calculated as the Euclidean distance. The submotion with the smallest difference obtained by this is classified and recognized as a submotion corresponding to the user's movement.

Third, the animation corresponding to the submotion is outputted and the next submotion is in standby state. We define submotions and animations in advance, and connect each submotion to other possible next submotions. The connected submotions are constructed in a graphical format with multiple continuous branches from one motion. When one submotion is completed, it waits for the user's next move and the above procedure is repeated.

4 Other Systems

The fighting game cannot be played by using the submotion system alone. In order to make the game playable, additional systems are developed and implemented.

4.1 Damage System

Previous VR games use simple collision detection system, i.e. the user colliding with the opponent inflicts a predetermined amount of damage on the opponent. However, by using submotion system, not only can submotions be categorized from user's motions, but other factors such as speed and motion size can be obtained. Using these data, the appropriate damage can be calculated for each motion. Firstly, the factors related to the submotion that has been recognized as an attack are obtained. Then, damage is calculated based on the programmed equation and measured factors such as submotion's acceleration, power, and motion size. Here, the damage is calculated by using a modified version of the well-known kinetic energy equation. Based on the size of the overall motion, the mass m can be obtained by using the preconfigured values, and the acceleration a can be calculated from the measured values. By plugging in the values, damage can be calculated. By calculating the damage in such a way instead of using pre-programmed values, the user can be motivated to make dynamic motions.

$$a = \frac{V - V_0}{t}. \tag{1}$$

4.2 Attack System

The categorized attacks, in addition to having an in-game HP-reduction effect according to the submotion type and damage, must also provide a visual effect based on the damage inflicted. Normally, various animations must be created and programmed based on various factors such as damage amount and damage location. Rather than creating and programming the animations in advance, however, the proposed method

here is to apply IK on top of the base animation. Unity's frame-level game logic progression order is Update, Animation, and LateUpdate. The IK is designed to be operated at the LateUpdate step (Fig. 2).

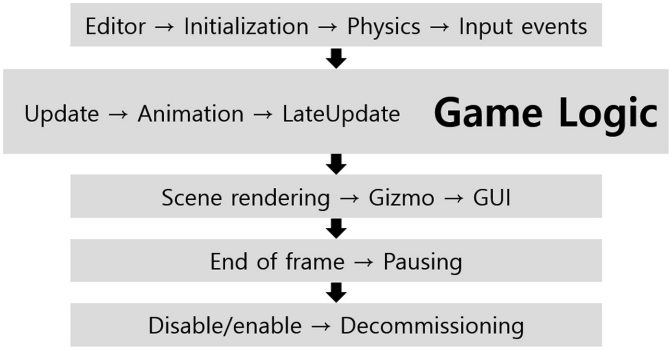


Fig. 2. Order of execution of event functions in Unity.

Another layer of animation is set to play on top of the base animation. First, two models are created: one in which the actual animation is executed, and an IK model. When a collision is detected in the animation model, a certain amount of force based on the damage inflicted is exerted on the corresponding part of the IK model, which moves it by a distance over a period of time and returns it to its original position. Since the animation is created based on the position on which the damage was inflicted and the amount of damage, its variation is practically infinite. Unique animations based on the damaged region and the amount of damage can be replaced by simple numerical properties for each region, such as weight values (Fig. 3).

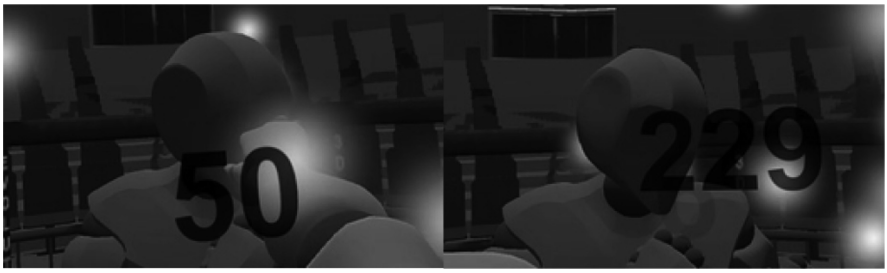


Fig. 3. Difference of damage and animation by kind of motion

4.3 Fighting Game System

In order to make the game more engaging, factors such as defense, evasion, and stamina are also added. The game emphasizes the importance of defense. Player can take advantage of the delay after an opponent guards itself, or attack regions that are

not defended by the opponent. Such counterattack and outmaneuvering attack provide a bonus damage, allowing for an advantage over the opponent.

This paper proposes a slow-motion effect in order to compensate for the lack of feedback in the game. By providing a slow-motion effect when the player attacks successfully or is attacked, more weight and feedback are given to the motions.

5 Conclusion, Further Plans

In this paper, we designed a full-body tracking VR fighting game and proposed a new motion recognition system named a submotion system to build this game. The submotion system recognizes the user's movement in a short unit, classifies it as a submotion corresponding to the movement, and outputs the corresponding submotion. The submotion system recognizes the user's movement and at the same time has freedom of motion through branching out various submotions. We suggested not only animation based on submotion but also game system elements that utilizes this such as damage and attacking systems.

This design provides new VR full-body tracking game contents satisfying the needs of VR users, additionally complements the shortcomings of the existing motion recognition methods and suggests motion recognition-based technology for future full tracking VR contents development. Various applications restricted in the existing motion recognition systems are possible with this system, which enables the user to give a purpose to each movement, to take more exaggerated and big motions, and to have more fun and immersion.

Acknowledgements. This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute for Information & communications Technology Promotion)" (2016-0-00017).

References

1. Niehorster, D.C., Li, L., Lappe, M.: The accuracy and precision of position and orientation tracking in the HTC vive virtual reality system for scientific research. Iperception, 8 (2017)
2. Shafer, D., Carbonara, C.P., Popova, L.: Controller Required? The Impact of Natural Mapping on Interactivity, Realism, Presence, and Enjoyment in Motion-Based Video Games. Presence Teleoperators Virtual Environ. **23**, 267–286 (2014)
3. Bowman, D.A., McMahan, R.P.: Virtual reality: how much immersion is enough? Computer **40**, 36–43 (2007)
4. Gobbetti, E., Scateni, R.: Virtual reality: Past, present and future. Stud. Health Technol. Inform. **58**, 3–20 (1998)

Smart Contract Based Academic Paper Review System

Min Jae Yoo and Yoojae Won(✉)

Department of Computer Science Engineering, Chungnam National University, Daejeon, South Korea
{vicerascal,yjwon}@cnu.ac.kr

Abstract. The centralized and opaque nature of the existing peer review systems and academic management practices of journals have led to various transparency and accountability issues. In this work, we propose a novel smart-contract-based review system that stores data using blockchain and web applications. The proposed method can realize transparent operation of the academic paper review process and improve the quality of research through comparisons of opinions, research communication, and reviewer compensation. In addition, the data generated during dissertation review processes can be accessed transparently, while ensuring security of the generated data.

Keywords: Smart contract · Blockchain · Review of academic papers · Academic communication · Transparency

1 Introduction

Since the advent of the academic research market, large publishing companies have been using distribution channels that have been built over several years as market forces, thereby forming a centralized ecosystem. This led to a closed review system for academic papers that was highly inefficient. Lack of transparency in reviews may result in problems with the peer review system. For example, in September 2015, Springer withdrew approximately 60 papers published under peer-review conditions owing to multiple forged identities that were used to manipulate the journal's online submission system [1]. In general, all academic journals use their own systems for anonymous assessment and self-examination [2]. The peer review system is inefficient owing to the unavailability of appropriate reviewers, which is exacerbated by the fact that reviewers are underpaid and their efforts are not recognized officially [3]. Thus, providing appropriate incentives can facilitate timely completion of the review process [4].

In this paper, we propose a novel academic paper review system using a smart contract to address the abovementioned problem. The proposed model utilizes a blockchain-driven infrastructure. The Ethereum platform provides services through a web application and smart contract, thereby ensuring security and transparency. By inculcating transparency into the peer review process, the quality of research can be improved by comparing opinions and enabling research communication and transparent compensation management. In addition, the system can be made robust against data forgery during the dissertation review process [5].

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 259–264, 2020.
https://doi.org/10.1007/978-981-13-9341-9_44

2 Related Studies

2.1 Web-Based Manuscript System

In the existing online manuscript submission system used by several academic journals, the screening process is opaque. It involves the researcher having to submit his/her manuscript and wait several months before knowing if the manuscript has been accepted or has been returned for revisions. During this period, the researcher is unaware about the progress of the peer review. Although the editorial structures may vary across journals, in general, they comprise editors and editorial committees, and the journal entry process includes many field processes. A typical review process for a dissertation is illustrated in Fig. 1; the stages include advance processing, e-mail collection, selection of reviewers, addition of reviewers, sending and waiting for manuscripts, and waiting periods for review.

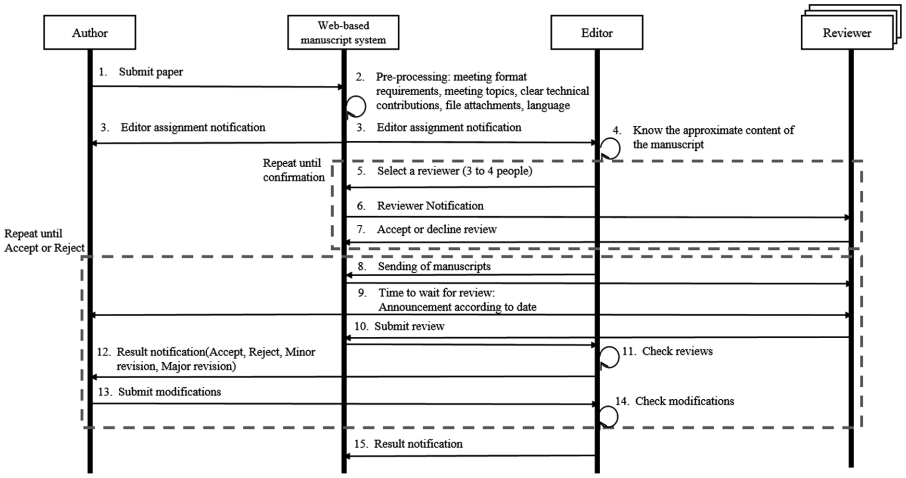


Fig. 1. Review process in web-based manuscript system

2.2 Smart Contract

A smart contract is a code that implements, enforces, and executes a contract. The concept of smart contracts was first realized in the Ethereum blockchain. It performs the role of an aggregator, enabling coordination between untrustworthy users and performing complex functions through blockchain technology. It has a user interface and provides better security than the traditional contract method with the logic of general contract clauses.

The smart contact on Ethereum refers to bytecode operating in the Ethereum virtual machine (EVM) and is run using the Solidity programming language [6]. Smart contacts on Ethereum consist of program codes, storage files, and account balances; the program code is fixed when the contract is created and cannot be changed [7]. Figure 2 illustrates the lifecycle of a smart contact.

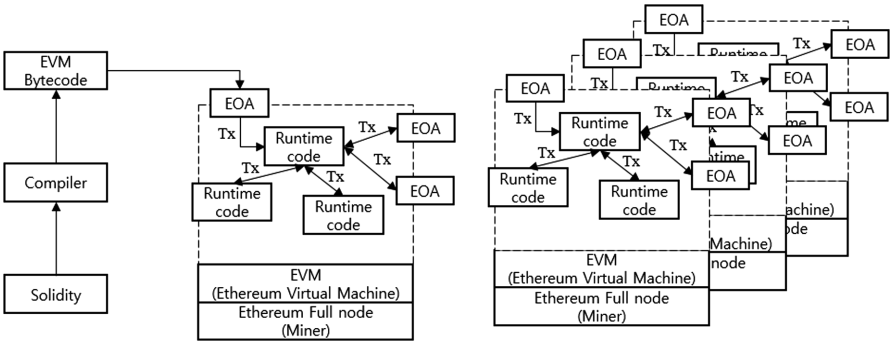


Fig. 2. Lifecycle of a smart contract

3 Academic Paper Review System Using Smart Contract

The proposed academic paper review system utilizes the smart contract system. In the proposed system, tokens are generated and distributed according to the contributions of the participants in the system. The distributed tokens are then used to register for paper review; the token generation, distribution, and destruction processes are automated and transparently disclosed through smart contracts. The proposed system utilizes the Ethereum platform to provide services through web applications and smart contacts. Thus, a more transparent and secure peer review system can be realized. Moreover, the proposed system allows review requestors to receive an additional number of unspecified opinions to accommodate various opinions on the study results, thereby enhancing the results of the peer review process. Aided by the interactive and transparent review process in the discussion space with general examiners, general users can review contentious findings by comparing their opinions with others and improve the quality of their research.

3.1 Suggested Structure of Academic Paper Review System

Using the smart contract system, the review requestor can contribute to the peer review process, and the reviewers and journal administrators can be automatically notified about the progress of the review. In addition, the general user can easily evaluate and discuss the findings of studies using author comments. The academy administrator manages the data through the blockchain server and web server and distributes tokens according to the contributions. The structure of the proposed academic paper review system is shown in Fig. 3.

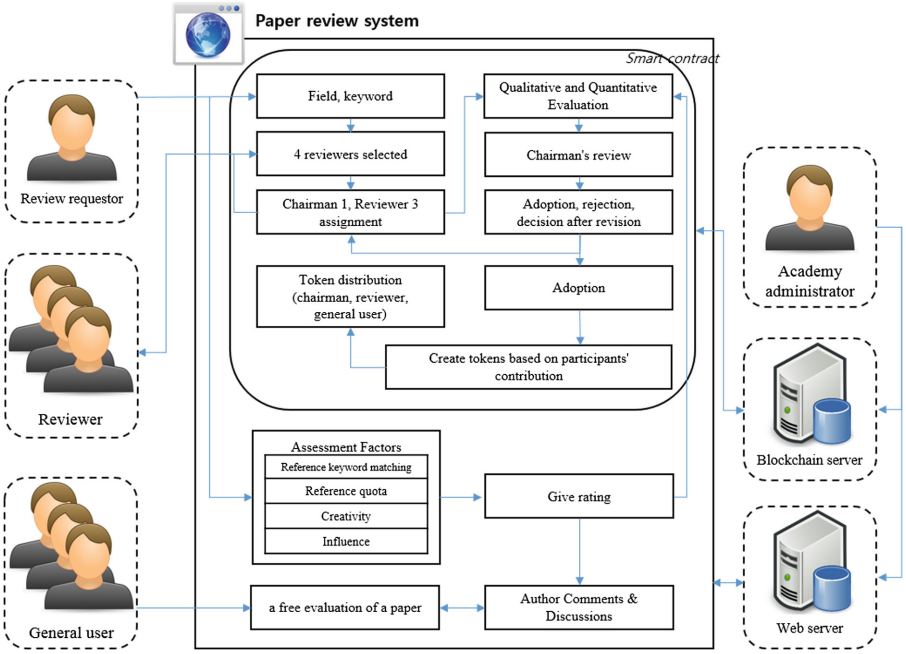


Fig. 3. Structure of academic paper review system designed using smart contract

3.2 Flowchart and Smart Contract Field

Figure 4 illustrates the flowchart of the actions of review requestors, reviewers, general users, and academy administrators. The action is performed as indicated in the flowchart, and the fields for the action are stored in the smart contract. One can send notifications to stakeholders when exiting or act through a smart contract.

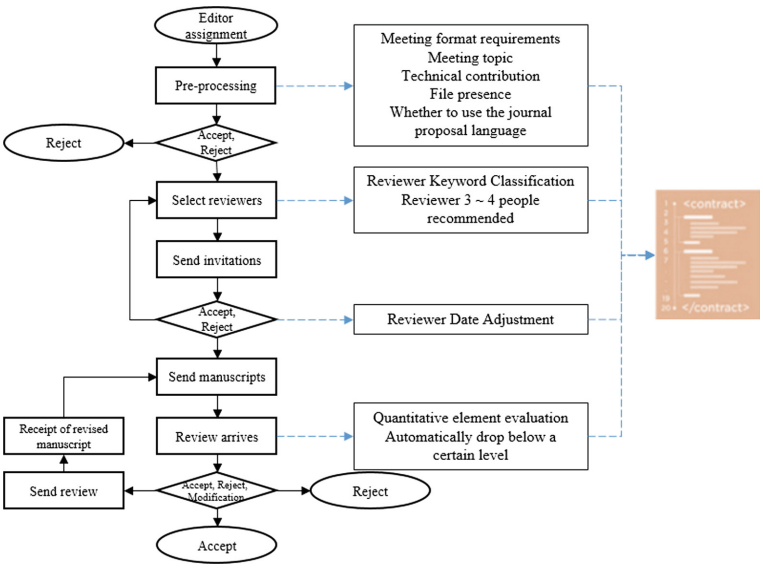


Fig. 4. Flowchart and smart contract field in the proposed system

3.3 System Scenarios Involving Smart Contracts for Academic Paper Review

Different system scenarios are illustrated in Fig. 5. The proposed system utilizes the process as applied in existing web-based manuscript systems, thereby allowing users to operate the system without discomfort. In addition, during this process, automation through smart contracts renders the system more convenient than the traditional system.

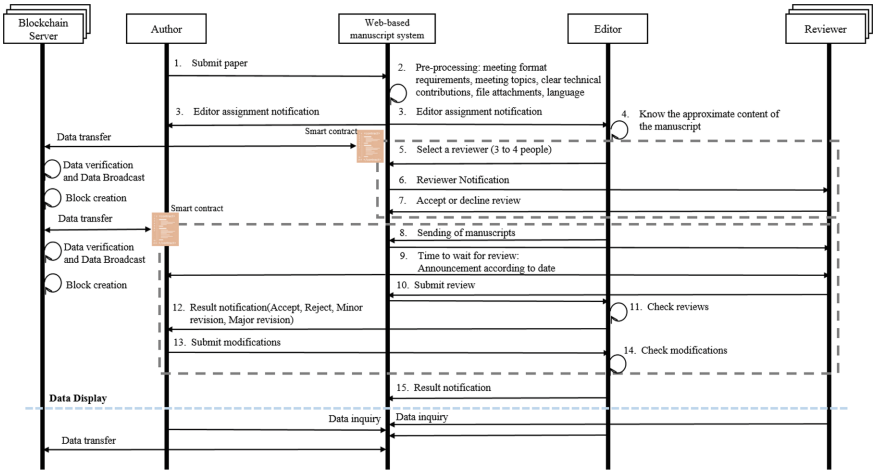


Fig. 5. Scenario involving the use of smart contracts for academic paper review

4 Conclusion

In this paper, we propose a smart contract-based peer review system to address the issues pertaining to transparency in the existing peer review process. Moreover, the proposed system also addresses the lack of adequate compensation for peer reviewers and solves the problem pertaining to the lack of diversity in opinions when assessed by a small number of reviewers. The proposed system can realize transparency in the academic paper screening process and improve the quality of research by enabling comparison of opinions, research communication, and using transparent compensation management. In addition, the overall transparency and security of the data generated during the thesis review process is enhanced.

In our future work, we will evaluate the performance of the proposed system using smart contracts to complement or replace the existing systems. In addition, we will experimentally test the hypothesis that the quality of paper review increases with the provision of incentives for reviewers. Further, we will motivate the participants by designing more detailed token economies that can use the tokens developed in this study.

Acknowledgments. This research was supported by the Ministry of Science and ICT, Korea, under the Information Technology Research Center support program (IITP-2018-2016-0-00304) supervised by the Institute for Information & communications Technology Promotion.

References

1. Kulkarni, S.: What causes peer review scams and how can they be prevented? Learn. Publish. **29**(3), 211–213 (2016)
2. Tite, L., Schroter, S.: Why do peer reviewers decline to review? A survey. J. Epidemol. Commun. Health **61**(1), 9–12 (2007)
3. Breuning, M., Backstrom, J., Brannon, J., Gross, B.I., Widmeier, M.: Reviewer fatigue? Why scholars decline to review their peers' work. Polit. Sci. Polit. **48**(4), 595–600 (2015)
4. Merrill, E.: Reviewer overload and what can we do about it. J. Wildl. Manag. **78**(6), 961–962 (2014)
5. Rodrigues, B., Thomas, B., Burkhard, S.: Enabling a cooperative, multidomain DDoS defense by a blockchain signaling system (BloSS). Semantic Scholar (2017)
6. Solidity. https://solidity.readthedocs.io/en/v0.4.25/. Accessed 24 Oct 2018
7. Delmolino, K., Arnett, M., Kosba, A., Miller, A., Shi, E.: Step by step towards creating a safe smart contract: Lessons and insights from a cryptocurrency lab. In: Clark, J., Meiklejohn, S., Ryan, P., Wallach, D., Brenner, M., Rohloff, K. (eds.) International Conference on Financial Cryptography and Data Security, vol. 9604, pp. 79–94. Springer, Berlin (2016)

Design of Image Generation System for DCGAN Based Picture Book Text

JaeHyeon Cho and Nammee Moon(✉)

Division of Computer and Information Engineering, Hoseo University, Asan, South Korea
nammee.moon@gmail.com

Abstract. When a picture book is photographed with a smart device, the text is analyzed for meaning and associated images are created. Image creation is the first step in learning DCGAN using class lists and images. In this study, DCGAN was trained with 11 classes and images of 1688 bears, which were collected by ImageNet for design. The second step is to shoot the image and text of the picture book on a smart device, and convert the text part of the shot image into a system readable character. We use the morpheme analyzer to classify nouns and verbs in text, and Discriminator learn to recognize the classified parts of speech as latent vectors of images. The third step is to create an associated image in the text. In the picture book, take the text of the part without the image and extract nouns and verbs. The extracted parts of speech and the learned latent vector are used as Generator parameters to generate images associated with the text.

Keywords: OCR · KoNLPy · DCGAN · Image generation

1 Introduction

Most early childhood institutions have a book area or language area that allows children to access various language activities and books. In this area, the infant reads the picture book himself or the teacher reads the picture book. Picture book reading activities are important activities in early childhood education institutions [1]. But a few years ago, digital picture books from tablet PCs and smartphones started to take its place [2]. Discussions with academics have begun claiming smart devices will have an impact on the life and development of infants. In fact, A research reported 90.3% of infants under the age of 7 years have used smartphones and 21.9% have used tablet PCs and other research reported that only 5.7% of infants are not using it compared to 95% of infants aged 4–6 using smart devices etc., it can be seen that smart devices have become familiar media for many infants in Korea. Thus, research on ways to utilize smart devices that have already been exposed to infants is already being carried out in both special education and general education. In terms of educational effect, infants have been found to utilize synesthesia through internet connection of smart devices and to communicate in various languages such as voice, video, and text, and to improve their creativity through app-based language activities. Research has also shown that smart devices can enhance young children's scientific thinking and problem-solving skills [3].

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 265–270, 2020.
https://doi.org/10.1007/978-981-13-9341-9_46

This study will be useful to utilize smart devices as educational media for young children. It takes images and texts from a picture book with a smart device, and extracts the text in a form that can be analyzed using OCR. Using the KoNLPy morpheme, the extracted text is classified into nouns and verbs, and learned by matching objects, behaviors, and images using Discriminator. In the picture book, extract only the text without the image, and We designate the noun and verb of the photographed text as the latent vector of the generator. The learned generator generates a similar image for the text.

2 Related Work

2.1 OCR

Fig. 1. Reading text from an image using OCR

OCR (Optical Character Recognition) is an image scanner that acquires images of human-written or machine-printed characters and converts them into machine-readable characters. OCR is a software that transforms the typeface of a document that can be obtained by scanning an image into a computer-editable character code and the like, and OCR is started as a research field of artificial intelligence or machine vision. Optical character recognition using optical technologies such as mirrors and lenses and digital character recognition by scanners and algorithms have been regarded as different areas, but now the term optical character recognition has been considered to include digital character recognition. The initial system required "Training", which meant reading a sample of the font in advance to read a font, but now most fonts can be converted to a higher probability. On some systems, you can create an output file in the same document format as a word processor file that closely matches it in the imported image. Among them (Fig. 1 Reading text from an image using OCR) Even if the image contains non-document parts like an image, it can be correctly recognized [4, 5].

2.2 KoNLPy

KoNLPy (Korean NLP in python) is a Python package for Korean information processing and is suitable for students who are just beginning to learn natural language

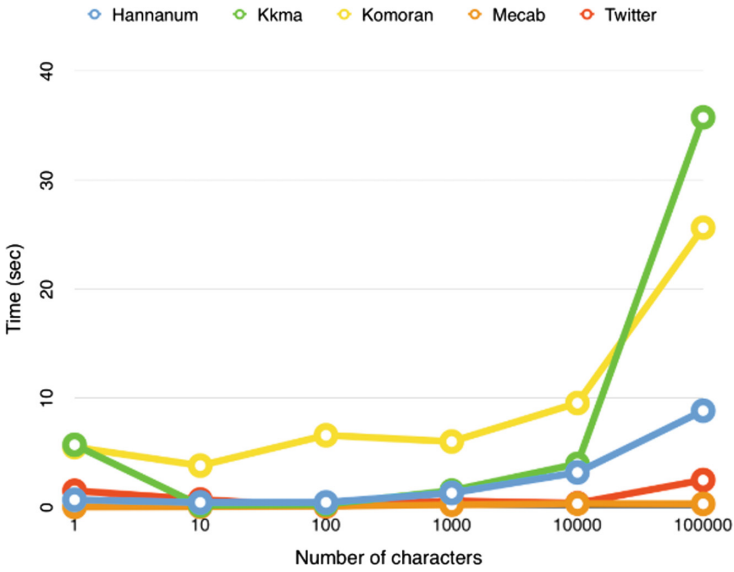


Fig. 2. KoNLPy morpheme speed comparison

processing or researchers who want to use natural language processing for research purposes. The KoNLPy morpheme analyzer is Kkma, Hannanum, Mecab, Okt, Twitter, etc. It can be used according to the user's situation such as loading time, execution time, spacing algorithm, word meaning and peripheral relation, word processing not included in dictionary [6].

2.3 DCGAN

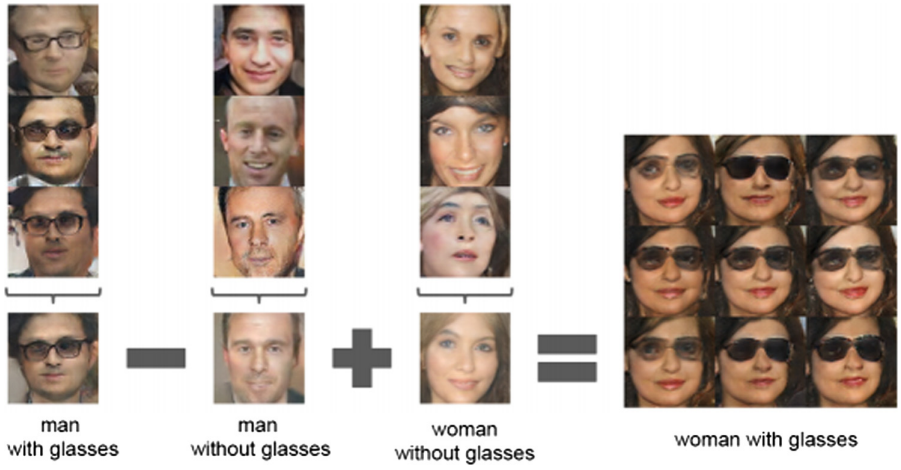


Fig. 3. Image generation using DCGAN's z operation

DCGAN (Deep Convolutional Generative Adversarial Network) is one of the deep-learning algorithms designed to understand images and find new images for them. For example, Remove the "male" learning result from the "glasses wearing man" learning result and add the "female" learning result to create the "Woman wearing glasses" as shown in (Fig. 3 Image generation using DCGAN's z operation) [7, 8]. Since the generative adversarial network (GAN) has solved the problem of minimax or saddle problem, it has a Convolutional GAN structure that can be learned steadily in most situations, and since the generator learned by DCGAN is capable of vector arithmetic operation, you can do sample generation at level [9]. In addition, DCGAN visualizes the filters learned, shows that certain filters have learned specific objects in the image, and the learned Discriminator has comparable image classification performance compared to other non-image learning algorithms [10]. A new method called z-arithmetic is also presented for the learned results by using the Gaussian noise vector as the noise input Z [11].

3 Method

The system of this study has three steps as shown in Fig. 3 (system design flow chart). First, the class improves accuracy with sorted data sets. The first step is to improve the accuracy of the classified data set. The most direct way to train DCGAN is to judge the text and image pairs together and train the discriminator to judge the pairs as real or fake.

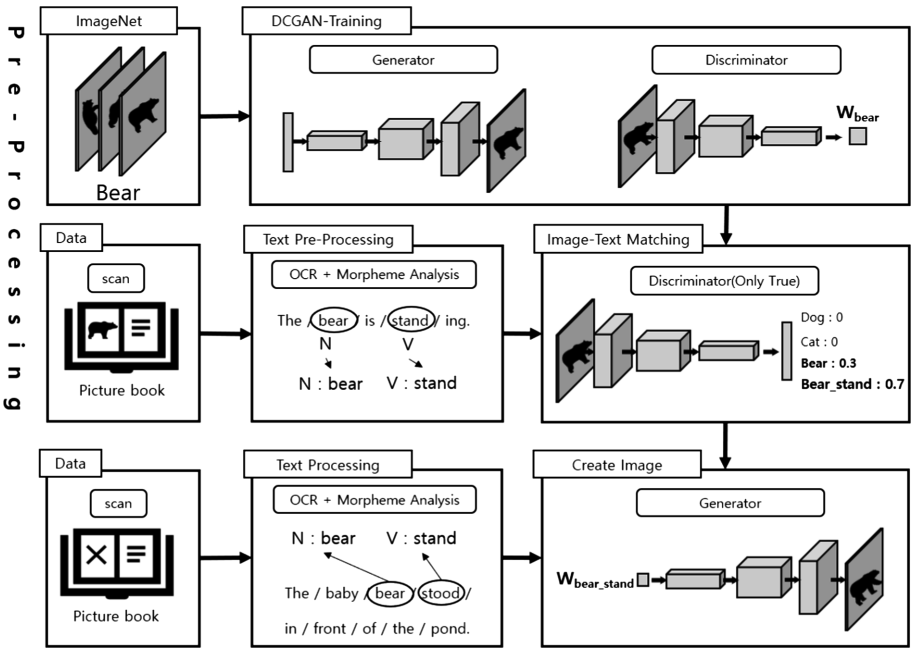


Fig. 4. System design flow chart

The data set uses 1688 bears images classified in 11 classes in ImageNet. We use a training image size of 64 × 64 × 3, a base learning rate of 0.0002, and the generator noise is sampled in a 100-dimensional unit normal distribution, training a mini batch size of 64 and epochs of 600. The Generator and the Discriminator alternate between the Generator and the discriminator network in order to compete and learn (Fig. 4).

The second step is to learn how to match images and text in a picture book. We take text and images from a picture book, and extract text using Tesseract, one of the OCR algorithms. In this study, Hannanum is used as KoNLPy morphological analyzer. (Figure 1 KoNLPy morpheme analyzer speed comparison), Hannanum has a relatively short loading time or execution time compared to other analyzers in analyzing many words, and a strength to fine-tune word processing that is not included in the dictionary have. The faster Twitter than Hannanum is currently stopped updating at KoNLPy, and Mecab does not support the Windows operating system. Images are classified into nouns and verbs. When Discriminator input a picture book image, it always learns 'True' (Fig. 2).

The final step is to create an image from the text. We extract text regions from the same picture book. As in the second step, we extract text using Tesseract and classify nouns and verbs using KoNLPy. A new image is generated by inputting the classified parts of speech as a latent vector of the generator [12].

4 Conclusion

This study will be useful to utilize smart devices as educational media for young children. In this paper, we propose a system design that learns text-image weights and generates images from text. The main factors that increase the accuracy of Tesseract, the OCR algorithm in the process of collecting data from picture books, are font size, illumination, and camera resolution.

Future studies will improve the accuracy of the OCR by adding pre- and post-processing of the shot, and if the training is sufficiently advanced, the image will be generated from the fairy tale with only text-free text. By gradually expanding the field of application and creating images in school textbooks, will be able to help the children, from students to students, in areas that they did not understand.

Acknowledgments. This work has supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2017R1A2B4008886).

References

1. Kim, I.-T., Yoo, K.-J.: Effects of augmented reality picture book on the language expression and flow of young children's in picture book reading activities. J. Korea Open Assoc. Early Child. Educ. **23**(1), 83–109 (2018)
2. Ryu, K.M., Kim, H.J., Kim, H.J., Lee, E.J.I., Heo, J.Y.: A development of interactive storybook with digital board and smart device. 한국HCI학회 학술대회, pp. 1179–1182 (2017)

3. Kim, Y., Park, H.: Study on the relation between young children's smart device immersion tendency and their playfulness. Early Child. Educ. Res. Rev. **20**(4), 337–353 (2016)
4. Lee, G.-C., Yoo, J.: Development an Android based OCR Application for Hangul Food Menu. J. Korea Inst. Inf. Commun. Eng. **21**(5), 951–959 (2017)
5. Smith, R.: An overview of the tesseract OCR engine. In: Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), Parana, pp. 629–633 (2007)
6. Park, E.L., Cho, S.: KoNLPy: Korean natural language processing in Python. In: Proceedings of the 26th Annual Conference on Human & Cognitive Language Technology, pp. 133–136 (2014)
7. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv:1511.06434 (2015)
8. Han, Y., Kim, H.J.: Face morphing using generative adversarial networks. J. Digital Contents Soc. **19**(3), 435–443 (2018)
9. Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., Lee, H.: Generative adversarial text to image synthesis. In: Proceedings of the 33rd International Conference on Machine. Learning, New York, NY, USA, 2016, JMLR: W & CP, vol. 48 (2016)
10. Springenberg, J.T., Dosovitskiy, A., Brox, T., Riedmiller, M.: Striving for simplicity: the all convolutional net. In: Proceedings of the International Conference on Learning Representations, pp. 1–14, arXiv preprint arXiv:1412.6806 (2015)
11. Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., Lee, H.: Proceedings of The 33rd International Conference on Machine Learning, PMLR, vol. 48, pp. 1060–1069 (2016)
12. Triantafyllidou, D., Tefas, A.: Face detection based on deep convolutional neural networks exploiting incremental facial part learning. In: Proceeding of the International Conference on Pattern Recognition, pp. 3560–3565 (2016)
13. Miller, E.L., Huang, G., RoyChowdhury, A., Li, H., Hua, G.: Labeled faces in the wild: a survey. In: Kawulok, M., Celebi, M., Smolka, B. (eds.) Advances in Face Detection and Facial Image Analysis, pp. 189–248 (2016)

Real Time Message Process Framework for Efficient Multi Business Domain Routing

KyoEun Kim[1], DongBum Seo[2], You-Boo Jeon[3], Seong-Soo Han[1], Doo-Soon Park[3], and Chang-Sung Jeong[4(✉)]

[1] Visual Information Processing, Korea University, Seoul 02841, Korea
comsdaq@korea.ac.kr, treeline@bestks.co.kr
[2] Advanced Technology Center, BestKS Co. Ltd R&D Center, Seoul 08595, Korea
postsky0@korea.ac.kr
[3] Department of Computer Software Engineering, Soonchunhyang University, Asan-si, Chungcheongnam-do 31538, Korea
{jeonyb,parkds}@sch.ac.kr
[4] Department of Electrical Engineering, Korea University, Seoul 02841, Korea
csjeong@korea.ac.kr

Abstract. In Message Oriented Middleware are composed of semiautonomous, heterogeneous, and independently designed next generation framework. For described to achieve successful operation of such a system, the activities of the message controlling the subsystems have to be coordinated. In addition, each layer applications can require synchronizes on events relative to message availability. A common way to proceed is to implement polling strategies within the distributed components, which leads to an increase of the load in the overall system. This approach are than describe a different based on informing using oriented message queuing. The message queuing processing of this technology in the multi business domain distributed system has been exercised.

Keywords: Message queuing · Distributed ubiquitous middleware · Context framework and middleware · Multi domain application · Message oriented middleware

1 Introduction

We can understand the challenges of real-time message processing and the fault of some of the existing solutions. In an effective multi-business domain, routing technology plays an important role in strategically determining data analysis and processing technologies. These technologies enable users to access massive business message using a real-time message processing framework to visualize results and provide lossless professional functionality. The real-time messaging framework is limited by the underlying message management infrastructure, and the messages must be preserved and recovered without loss when the system makes real-time requests. In particular, there is a need for a system that preserves expertise without loss when routing across multiple domains. These components can be implemented by adding separate

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 271–278, 2020.
https://doi.org/10.1007/978-981-13-9341-9_47

functions to handle large amounts of messages. Separately storing, analyzing, and processing messages in real time is a different scenario from lossless message handling. In multi-domain routing, separate message processing technology requires a design that realizes real-time processing and lossless transmission by implementing a separate technology in a large amount of messages processed in real time and implementing it in architecture. Message control is a specialized agent in ICON_UBM (Icon for Unix Base Middleware) field, message and institutional message information memory to automatically convert external specializations to internal specialties of ICON_UBM and internal specializations of ICON_UBM to external specializations. It provides fast, accurate and convenient service for maintenance. This paper explained the related research to presents a design for an UNIX based middleware was where the need a tool arises that can handle real time message processing and analytics.

2 ICON_UBM Architecture

The Message Oriented Middleware (MOM) uses messaging providers to mediate messaging operations. The basic elements of the ICON_UBM system was client, message, and ICON_UBM providers, which include APIs and management tools. ICON_UBM providers can use different structures to route and forward messages.

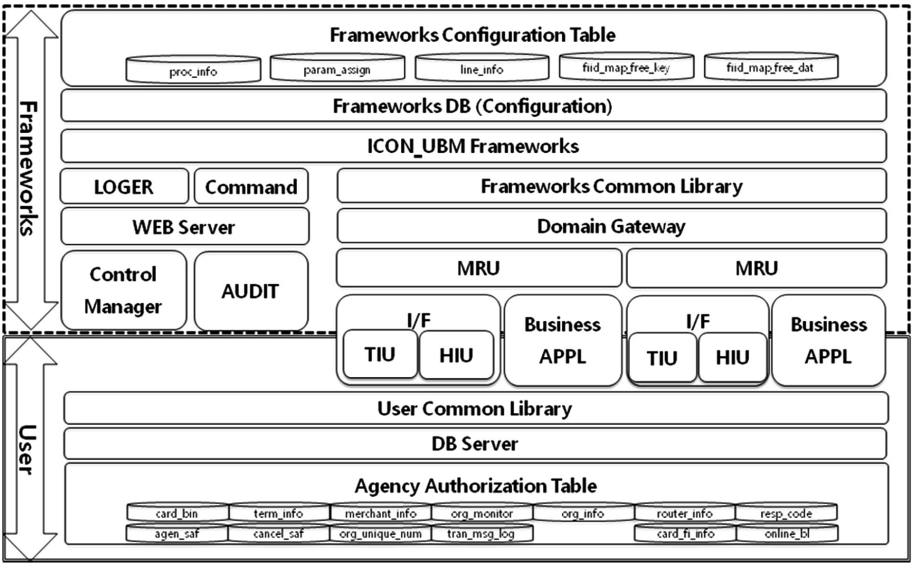


Fig. 1. System Architecture of ICON_UBM

ICON_UBM providers could used a centralized message server or deploy routing and forwarding capabilities to each client computer. Some ICON_UBM products used a combination of these two methods. Clients invoke APIs through the ICON_UBM

system to send messages to destinations managed by the provider. As a result, client applications could effectively escape from all issues except those related to sending, receiving, and processing messages. Our analysis of these architecture as shown in Fig. 1. The ICON_UBM frameworks supports all functions that it could provide internally with each layer operating. The ICON UBM was at the heart of the framework, effectively a large number of simultaneous queries against a live message set [1].

2.1 Frameworks Layer

The Monitoring of logs generated by each layer of ICON_UBM system was performed by tool of web environment for integrated log management. The Memory Control Layer (MCL) is ICON_UBM service agent that delivers messages and events between service processes in the system. It was divided into synchronous and asynchronous according to the service method and provides queuing service by store and forwarding method. Supports emergency message processing by using timeout management and message priority function. The Frameworks Message Control Layer (FMCL) offers was a specialized agent in ICON_UBM that loads field, message and information into the memory to automatically convert the internal specialization of ICON_UBM into the internal specialization of ICON_UBM and the internal specialization of ICON_UBM to external specialists. It was fast, accurate and convenient. The Frameworks Communication Control Layer (FCCL) provided communication adapter functions such as TCP/IP Adapter and XML Adapter, so that it could easily implement the linkage function with external organizations or internal systems in any communication environment. In addition, a new system linkage function could apply it immediately by parameter setting and change only by using administrator function. The Service Control Layer (SCL) admin tool of the web environment was a system that collectively manages the environment file information required by each layer, and it focuses on the convenience of easy access and maintenance by the administrator.

2.2 Terminal Interface Unit

The Terminal Interface Unit (TIU) was analyzes the card authorization or cancellation telegram from various terminals and normalizes them and transfers them to the router. Line management module manages the state of terminal equipments and gives flexibility to easily cope with new equipments without adding big modification of whole system even when adding new communication methods such as new types of terminals or mobile to message supported encoding or decoding in a certain way. Timer was activated after receiving all text messages, and if the result was not received within a certain time, automatic ‘FAIL’ processing. Monitor Agent (MA) was an application and transmits information such as status information, professional processing information, and log in conjunction with monitor unit to logging all the specializations in real time as needed.

2.3 Message Router Unit

The message transferred from the TIU was distributed to the corresponding application by the Message Router Unit (MRU) and is transmitted to the other domain through the Multi-Domain routing Gateway (MDGW) in order to call the task of the other domain. The App and Scheduler manages multiple application processes for the same job and performs distributed processing. By managing each application-specific queue, it was possible to process without overloading or loss of expertise in case of a specific application overload or failure. In addition, Application hot swap was supported, so that it is possible to deal with failures and upgrades of specific applications without loss of expertise. MA was an application it responsible for transmitting status information, professional processing information, and log information in conjunction with the monitor unit. Auditor provided a function to check detail log in real time to verify suitability of professional processing.

2.4 Multi-domain Routing Gateway

The Each domain has its own Multi Domain routing Gateway (MDGW), and MDGW was only responsible for invoking tasks of other domains. The result of the processing of one MDGW request was that the MDGW of the domain that processed the task calls the task of the original domain and returns the result. In this way, there was no need to wait for MDGW to perform processing of another domain, and the processing procedure was simplified, thereby eliminating the bottleneck due to Gateway and efficiently managing complex tasks among multiple domains.

2.5 Host Interface Unit

The Host Interface Unit (HIU) was a process for processing specialist information associated with a credit card company or other financial settlement system and works in real time with the system of the financial institution. It was based on processing the specialization that meets the standard of ISO8583 which was the international standard format. In addition, it was composed of modules capable of processing special specialization according to the characteristics of each financial institution. Monitor Agent was an Application. It was responsible for transmitting status information, professional processing information, and log information in conjunction with the monitor unit.

2.6 Application Monitor Unit

The Application Monitor Unit (AMU) monitors online applications and batch applications that make up the framework and approval system, and plays the role of replaying in real time if it was down. It also acts as an agent for management and transmits information of each unit of framework to EMS (Enterprise Management System). It manages or controls the authority of users who could access Framework through Authentication Management Module (AMM). It has a separate CUI (Command line User Interface) and provided emergency fault handling function.

3 Heuristic Demand Model

The heuristic demand model has issues related to the interoperability, reliability, security, scalability, and performance of the various MOM systems resolved within the MOM system implementation [2] and [3]. There are some drawbacks to this approach, and we have developed a system structure called ICON_UBM. ICON_UBM writes both questions and information to other CPU centers like multi-domain. For optimal configuration, the task that the host chooses when the server is set up for a given station is somewhat complicated, depending on the number of operations on the other station. In the heuristic Demand model condition [4] and [5], the server services point-in-time 1, which processes Q working or queue 1 occurs first. Thus, our heuristic Demand model condition is switched from queue 1 to queue 2 whenever $P_1 = 0$ and $P_2 > 0$, or $P_2 = Q$. The server is idle in Queue 1 when $\sum_{l=1}^{h} P_i = 0$. The optimal condition for switching from queue 1 to queue 2 contains a more complex transition curve. Each time we repeat this process we count the expected outcome of the condition specified by the queue value and the vector value $z = (z_2, z_3, \ldots, z_{N-1})$. So that when the server switches to queue 2, the number of queue 2 operations is between 1 and queue. The expected per-job condition when a package is processed indicates the probability of processing the size placement of the server in queue 1 before switching to queue 2 when using the conditions defined by queue. The amount of time that a working waits on translation 1 before the package that the job belongs to begins. We define n_{dl}^{Q} represent the number of package processed at translation d when the server switches to translation 2 after processing l jobs at translation 1. We define $c_l(Q, z)$ represent the expected sum of the holding costs happened by a package of size l from the time the first unit in the package is processed at queue 1 until the last unit is closed at queue N. This cost can in a hurry be computed truly using a recursive sequence. The expected cost per working when a batch of size l is processed is than given by $c_l(Q, z)/l$. Let $p_l(Q, z)$ represent the probability that the server processes a package of size l at queue 1 before switching to queue 2 when using the conditions defined by Q and z. Let $w(Q, z)$ represent the expected amounted of time that a working waits at translation 1 before the package that it belongs to starts being served. Then, our valuation of the average cost per working under condition (Q, z) is given by

$$Q, z = h_1 w(Q, z) + \sum_{l=1}^{Q} pl(Q, z) \frac{cz(Q, z)}{l} \tag{1}$$

Therefore, we necessary to get valuation of $p_l(Q, z)$ and $w(Q, z)$ in order to compute Q, z. The valuation for $w(Q, z)$ is got by impertinent that package of size Q get processed each time. Factually, we transform the original system with one where working arrive only in package of Q. Using our fraction sequence, the amount of time that the server would spend on processing a package of j through all the phase has average

$$m_l = \sum_{i=1}^{h} (h_{dl}^{Q} s_d + lb_d), \tag{2}$$

and variance

$$v_l = \sum_{d=1}^{h} (h_{dl}^{Q} Var[S_d] + lVar[B_d]). \tag{3}$$

After this, the system where working arrive in package of Q and are served in package of Q is a $GI/G/1$ queue with processing time for a whole package with average m_Q and variance v_Q. Because in the system, the inter arrival time is false to be exponential with rate λ, in this system the inter arrival time of a package of size Q is reference [6]. Let $\rho_Q = \lambda m_Q/Q, C_{aQ}^2 = 1/Q$ and $C_{sQ}^2 = v_Q/m_q^2$. Also C_a^2 is variance divided by the square of the average to an inter arrival time and C_s^2 is service time distribution with this heuristic demand model thus depends only on the basic parameter for multi domain routing. Than using the waiting time closeness for a $GI/G/1$ queue [7] our valuation for the waiting time is

$$w(Q, z) = \frac{Q\rho_Q^2(C_{aQ}^2 + C_{sQ}^2)}{2\lambda(1 - \rho Q)} \tag{4}$$

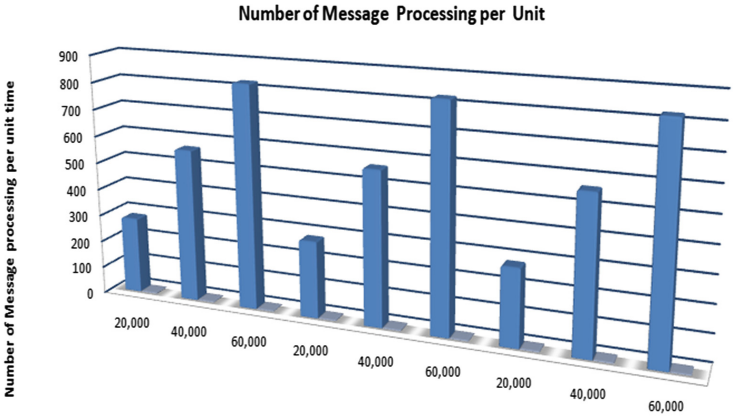
A real-time message processing framework was used in order of being sorted as part of a memory table using various strings, filtering out the unique range of messages.

4 Experiment

In order to test the number of message processing per unit time, we measure the number of processing that the TIU received from the simulator was sent from TIU_-Send, processed by MRU, passed through HIU, and returned to TIU_Receive after MRU with more than 800 targets per second do. The test objective for message throughput was to measure the throughput of the specialist sent from TIU_Send after passing through the HIU and then returning to the TIU_Receive after passing through the MRU and target was more than 90%. The number of types of messages that could be processed was three types of messages transmitted from TIU_Sender received by the MRU, and after passing through the HIU, it was checked whether the MRU returned to the TIU_Receive. For test, the implementation was accomplished on 2 PCs as clients, and clusters (CPU: Inter® Core™ i5-7200U CPU @ 2.50 GHz RAM: 8 GB Dual) with client software was Windows 10 Home x64, SecureCRT 5.2, Internet xploer11 and computer as servers (CPU: Xeon 2.4 GHz 4 core, Memory: 32G, Disk: SAS600G * Dual) with Server software was CentOS Linux release 7.2.1511, Linux version 3.10.0-327.el7.x86_64, JDK version 1.8.0_65, JDK Runtime Environment (build 1.8.0_65- b17), Apache-Tomcat/8.0.45, SQL-Lite 3.7.17, Logger, MRU, HIU, TIU, Log-Viewer [3] (Table 1, Fig. 2).

Table 1. ICON_UBM Frameworks experimental test bed condition

Condition	Exam Goal	Perform
Number of Message	800/second	1. Create Table
Message Throughput	More than 90%	2. Check basic data load and loaded data
Number of Message Type	3 type	3. Set MRU Configuration
		4. Set Message Logger Configuration
		5. Set Logger Configuration
		6. Identify program files



	20,000	40,000	60,000	20,000	40,000	60,000	20,000	40,000	60,000
Total time	283.18852	566.82380	827.19827	283.18985	566.86930	827.20970	283.20187	566.86930	827.17540
Process Type	3	3	3	3	3	3	3	3	3

Fig. 2. Simulation Value of number of message processing per unit time

5 Conclusion

The result from this proposal we have presented the real time process frameworks on developed through this study is a high-availability system that processes data in real time without loss of messages and supports the standardization of international standard ISO-8583, minimizing customization in system linked with. In addition, the findings from these analyses suggest that Multi-Routing is supported for specialized, so that a complex system can be divided into several systems, and a message can be transmitted and processed while minimizing bottlenecks. A series of tasks derived from one message can be executed by multiple systems and provides a distributed processing function that can be performed. The operating system framework can be flexibly extended to develop next-generation financial services as well as financial settlement services easy expansion to similar financial services. By integrating and managing data access through a financial settlement operating system framework into a single framework controlled by a framework rather than individual applications, it is possible to ensure the integrity of data and ensure stability as a financial system. In this regard,

the authors have designed various module configurations to use the framework. In addition, it is possible to provide a UI that can be easily used by customers without complicatedness. In general, the performance of program messages running at processing is about 100% lower than messages running in the ICON_UBM framework. We were able to get test results using our own lossless professional framework and have completed the development of a modular framework. This is an improved the implementation of message queuing technologies using the heuristic demand model to the ICON_UBM. Furthermore, message queuing proved to be efficient when utilize as an interface TRU, MRU, TIU from external agent.

Acknowledgments. This work was supported by the Advanced Technology Center funded by the Ministry of Trade, Industry and Energy (MOTIE, Korea). [Project Name: Developing financial settlement system and e-money trading system based on distributed ledger technology for developing countries; Project Number: 10077325] and This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2018-2014-1-00720) supervised by the IITP (Institute for Information & communications Technology Promotion).

References

1. Meester, L.E., Shanthikumar, J.G.: Concavity of the throughput of tandem queueing systems with finite buffer storage space. Adv. Appl. Probab. **22**(3), 764–767 (1990)
2. Patel, K., Sakaria, Y., Bhadane, C.: Real time data processing frameworks. Int. J. Data Min. Knowl. Manag. Process (IJDKP) **5**(5), 49–63 (2015)
3. Duenyas, I., Gupta, D., Olsen, T.L.: Control of a single-server tandem queueing system with setups. Oper. Res. **46**(2), 218–230 (1998)
4. Zayas-Cabán, G., Xie, J., Green, L.V., Lewis, M.E.: Dynamic control of a tandem system with abandonments. Queueing Syst. **84**(3–4), 279–293 (2016)
5. Armstrong, J.: Making reliable distributed systems in the presence of software errors. Diss. Mikroelektronik och informationsteknik (2003)
6. Whitt, W.: Approximations for the GI/G/m queue. Prod. Oper. Manag. **2**(2), 114–161 (1993)
7. Krämer, W., Langenbach-Belz, M.: Approximate formulae for the delay in the queueing system GI/G/1. In: Proceedings of the 8th International Teletraffic Congress, vol. 2, no. 3 (1976)

Malware Classification Using Machine Learning

JinSu Kang and Yoojae Won(✉)

Department of Computer Science Engineering, Chungnam National University, Daejeon, South Korea
{kangjsl798,yjwon}@cnu.ac.kr

Abstract. Recently, tools for generating malware have been spreading rapidly on the internet, making it easier for people without expertise to create malware. As a result, the number of malware variants is increasing quickly. To address this issue, it is crucial to classify malware quickly and accurately. However, malware variants are evolving to evade traditional malware-detecting methods based on signature pattern matching. To solve this problem, researches on detection of malware have been made in various fields. In the present study, we first propose a classification method to extract feature data from malware files that is applicable to machine learning, and then we classify malware through learning. Finally, we apply our classification method to sample data to evaluate performance and analyze the results.

Keywords: Computer security · Metamorphic · Polymorphic · Malware classification · Machine learning

1 Introduction

Recently, tools for generating malware have been spreading rapidly on the internet, making it easier for people without expertise to create malware. As a result, the number of new malware variants is increasing quickly. According to AV-Test malware trend report, the number of recent malware has been increasing, but the number of new malware has not changed much. Most of the recent malwares are variant of existing malware [1]. Most of the new malware use either a polymorphic method for compressing and encrypting existing codes or a metamorphic method for transforming files to detour signature pattern matching-based malware detection employed in obtaining traditional anti-virus solutions. Currently, various malware analysis methods are being studied in response to the generation of malware variants. In the present study, we propose a malware classification method using machine learning, for which attribute data applicable to learning is extracted from malware data. This is followed by an analysis of the obtained results.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 279–284, 2020.
https://doi.org/10.1007/978-981-13-9341-9_48

2 Training Data Set

To carry out this study, we use the Microsoft Malware Classification Challenge data set (BIG 2015) hosted by Kaggle in 2015 [2]. For each data set, approximately 10,000 samples are provided in byte and disassembly file format to both the learning and the test data. Each data set is contained in one of the classes listed in Table 1.

Table 1. BIG 2015 data set classes

Number	Family name	Type
1	Ramnit	Worm
2	Lollipop	Adware
3	Kelihos_ver3	Backdoor
4	Vundo	Trojan
5	Simda	Backdoor
6	Tracur	Trojan downloader
7	Kelihos_ver1	Backdoor
8	Obfuscator.ACY	Obfuscated malware
9	Gatak	Backdoor

3 Feature Extraction

This section discusses the feature data extracted that is to be used for machine learning. In the present study, we extract the section name data of sequences, API, image data, and instruction and assembly code as features to use for learning. Each feature is explained in the following subsections.

3.1 Sequences

Malware files can be classified by analyzing byte sequences as they maintain binary format. For sequence analysis, the N-gram method is widely used. In the N-gram method, a long string is divided into multiple strings of size N, and then a statistical technique is employed to analyze the pattern of each fragment. In this study, we set N to be 4 and extract 4-gram data in the binary files of the malware sample data to use for learning [3].

3.2 Application Programming Interface (API)

API information regarding where a specific file is being imported can be identified by PE file analysis. API information is related to actual performance for PE file execution; thus, it is a key indicator for malware analysis. However, many recent malware variants exploit obfuscation; hence, API information to be extracted from files is limited. This makes it difficult to use API information alone for learning and requires its use in

conjunction with other features [4, 5]. In the present study, we measure the number of API calls in the malware sample files and store it to use as learning data (Fig. 1).

Fig. 1. API call extraction.

3.3 Section Name

The PE file is configured with pre-defined section names, including .text, .data, .idata, .edata, .rdata, .rsrc, .tls and .reloc [6]. Each section includes a code to execute a program, a global variable, and DLL information for importing and exporting resource-related data. However, the PE file is changeable because file characteristics are not mandated. While normal files tend to maintain the format of section names, malware uses an arbitrary section name for its obfuscation. Accordingly, section names are extracted as a feature to be used for learning. In our research, we extract the section names existing in the learning data and count the number of each file's section lines to use for learning.

3.4 Instruction

The binary file is difficult to intuitively understand and thus becomes tricky to analyze. For more effective analysis, the assembly code of a file needs to be considered. In fact, it is in the assembly code that instruction information can be found [7]. The sequence can be analyzed using the N-gram method as one does in a byte file. In this study, we measure the frequency of instruction to use it as feature data.

3.5 Image Representation

Because malware variants are visually identical to existing malware via imaging, recent studies have carried out the imaging of malware and then employed image-treating algorithms such as convolution neural networks (CNNs). Figure 2 shows the method for imaging malware. Imaging is performed by transforming each byte in the malware binary file into an 8-bit vector and subsequently turning it into malware, which is represented by the greyscale image in Fig. 2, with values between 0 and 255 [8].

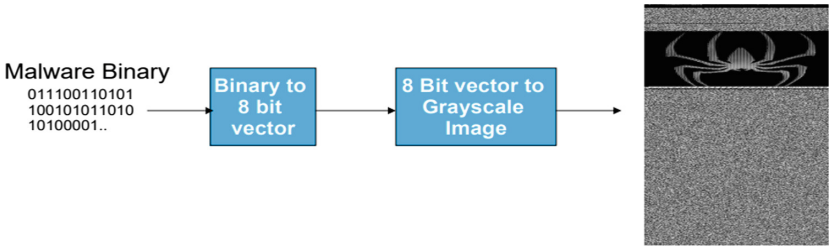


Fig. 2. Visualizing malware as an image.

Imaging is possible for both binary files and assembly files of malware sample data. In the present study, the images of the assembly files are clearer than those of the binary files, and thus they are used as features. Figure 3 shows the image pattern of each class after imaging malware learning data. In this study, we extract 8-bit vectors from the malware binary files and store them to use as learning data.

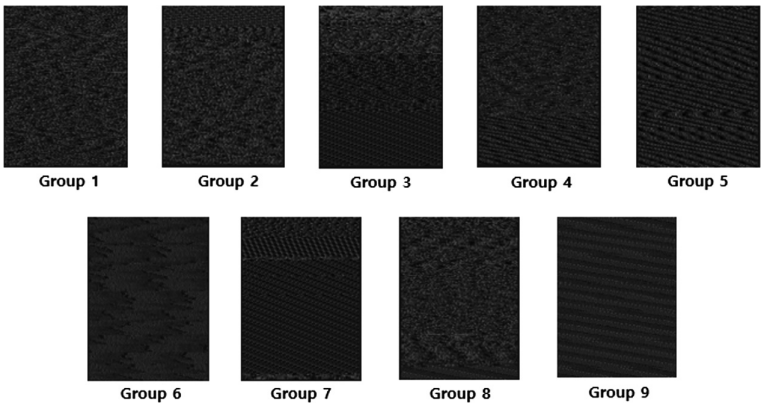


Fig. 3. Training a malware image.

4 Learning and Performance Evaluation

This section discusses the results of extracting the feature data discussed in Sect. 3 from malware sample data and employing machine learning. In the present study, we conduct learning using random forest, an ensemble technique-applied machine learning algorithm, followed by an assessment of the algorithm's performance. Random forest is an algorithm that creates multiple decision trees during training and performs classification and prediction of the mean value. In this study, 75% of the entire data is used as a training set for the experiment, and the remaining 25% is used as a test setup to assess performance. In addition, the out-of-bag (OOB) score is measured while applying the random forest to assess performance (Fig. 4).

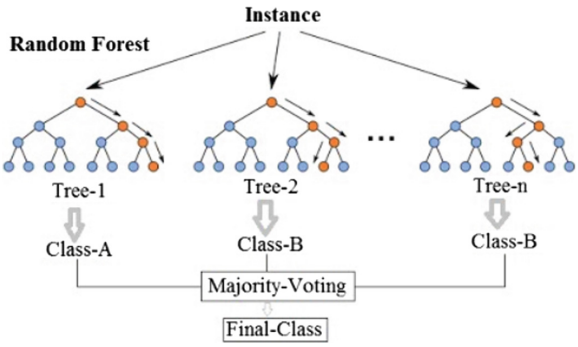


Fig. 4. Random forest simplified.

Table 2 shows measured performance. After training the model using the training set, the performance was measured using the test set, and the accuracy was measured as 99.8%. Likewise, we measured the performance using OOB data and achieved an accuracy of 99.5%.

Table 2. Evaluation result

Data	Accuracy
Test set	99.8%
Out of bag	99.5%

Figure 5 shows a confusion matrix. We have confirmed that most of the data are correctly classified.

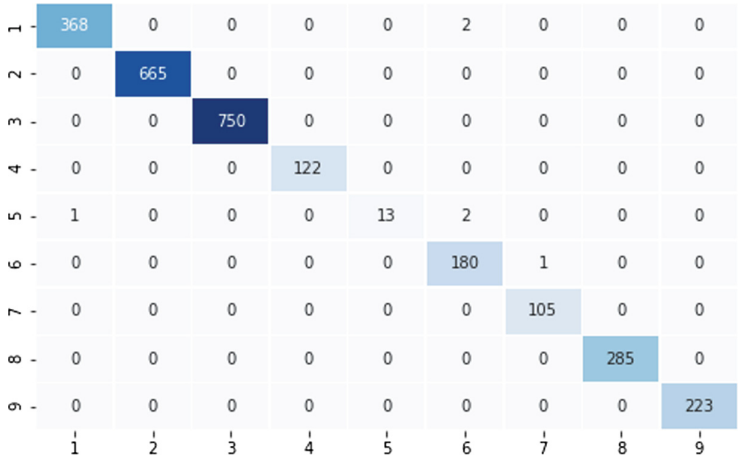


Fig. 5. Confusion matrix

5 Conclusion

Recently, many studies have been carried out to deal with the rapid emergence of malware variants. In the present study, we propose a method for classifying malware using machine learning and conduct related experiments. After performing the learning procedure using the Microsoft Malware Classification Challenge data set, we analyze its performance, and an accuracy of 99%. In future research, we will develop new methods for improving the performance of malware classification. In addition, we will explore how to develop a data set using both normal and malware files to analyze malware-detection performance.

Acknowledgments. This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2018-2016-0-00304) supervised by the IITP (Institute for Information & communications Technology Promotion).

References

1. AV-Test: Security Report 2016–2017 (2016)
2. https://www.kaggle.com/c/malware-classification
3. Santos, I., Penya, Y., Devesa, J., Bringas, P.: N-grams-based file signatures for malware detection. In: Proceedings of the 11th International Conference on Enterprise Information Systems (ICEIS), AIDSS, pp. 317–320 (2009)
4. Islam, R., Tian, R., Ba, L.M., Versteeg, S.: Classification malware based on integrated static and dynamic features. J. Netw. Comput. Appl. **36**, 646–656 (2013)
5. Ki, Y., Kim, E., Kim, H.: A novel approach to detect malware based on API call sequence analysis. Int. J. Distrib. Sens. Netw. **11**(6), 4 (2015)
6. Ahmadi, M., Ulyanov, D., Semenov, S., Trofimov, M., Giacinto, G.: Novel feature extraction, selection and fusion for effective malware family classification. In: Proceedings of the Sixth ACM Conference on Data and Application Security and Privacy (2016)
7. Santos, I., Brezo, F., Nieves, J., Penya, Y., Sanz, B., Laorden, C., Bringas, P.: Idea: opcode-sequence based malware detection. In: Engineering Secure Software and Systems. LNCS, vol. 5965, pp. 35–43 (2010)
8. Nataraj, L., Karthikeyan, S., Jacob, G., Manjunath, B.S.: Malware images: visualization and automatic classification. In: Proceedings of the 4th ACM Workshops on Security and Artificial Intelligence, pp. 21–30 (2011)

Operating System Fingerprint Recognition Using ICMP

Jinho Song, Yonggun Kim, and Yoojae Won(✉)

Department of Computer Science and Engineering,
Chungnam National University, Daejeon, South Korea
{thdwlsgh91, sasinsg625, yjwon}@cnu.ac.kr

Abstract. The operating system fingerprint is a factor that can help determine a target operating system and version through network scanning. There are two methods of discrimination: Internet Control Message (ICMP) and Transmission Control Protocol (TCP). In this study, we analyzed whether it is possible to categorize the operating system version (e.g., Windows 7, 8.1, 10) in a manner that the operating system can be determined using ICMP. Using ICMP, we could successfully classify the operating systems into Windows and Linux.

Keywords: ICMP · Operating system · OS fingerprint · Nmap · Wireshark · Network packets

1 Introduction

New types of network services are being continuously developed for widespread application of network technology, Many products such as home appliances, mobile phones, smart homes, wearable computers, etc. have appeared. However, with the increase in this development, the problem of security has become a major threat to the social infrastructure as a whole. The various strategies and methods of hacker attacks have been researched to improve the structure of the network security system and to minimize losses due to network problems. Data collection is the first step before a hacker attacks a particular system, and network searches are used as an important means of data collection. The purpose of the network scan is not limited to confirming that the connected system is alive or discovering the open port, but it is also used to collect operating system (OS) information and version (e.g., Windows 7, 8.1, 10). In this paper, we describe an analytical method that uses ICMP communication to distinguish the type and version of the OS used by the target host.

2 OS Fingerprint

OS fingerprint [1] is a technology to find the OS of the device to be identified by using various values according to the OS default setting, and it is possible to guess the OS through the response received using various commands to the device. The information used to guess the OS corresponds to the TCP/IP protocol such as Time to Live (TTL), Windows Size, Max Segment Size (MSS) and Packet Length. The OS fingerprint can be identified using active or passive fingerprint methods.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 285–290, 2020.
https://doi.org/10.1007/978-981-13-9341-9_49

2.1 Active Fingerprint Method

The Active Fingerprint method transmits network packets to the user device and analyzes the packet length, TCP/IP options, etc. in the response packet to guess the operating system. There is an advantage that the analysis speed is fast, but there is a disadvantage that the fact that the present device is detected comes to the user.

2.2 Passive Fingerprint Method

In the passive fingerprint method, a packet is not directly transmitted to the device, and it is instead identified in the received packet. It is different from the active method in that one cannot notice that the target device is being scanned because incoming packets are used; however, a disadvantage is that the analysis of incoming information requires a long time. In this work, active fingerprint ICMP communication is used to determine the OS.

3 Internet Control Message Protocol

The ICMP is a protocol used to notify a user of the problems that occur during processing of the TCP/IP and IP packets, and to perform other functions necessary for the IP layer, such as diagnosis. Because an ICMP message is processed as IP protocol data, it is encapsulated in the IP header and transmitted.

3.1 ICMP Message Format

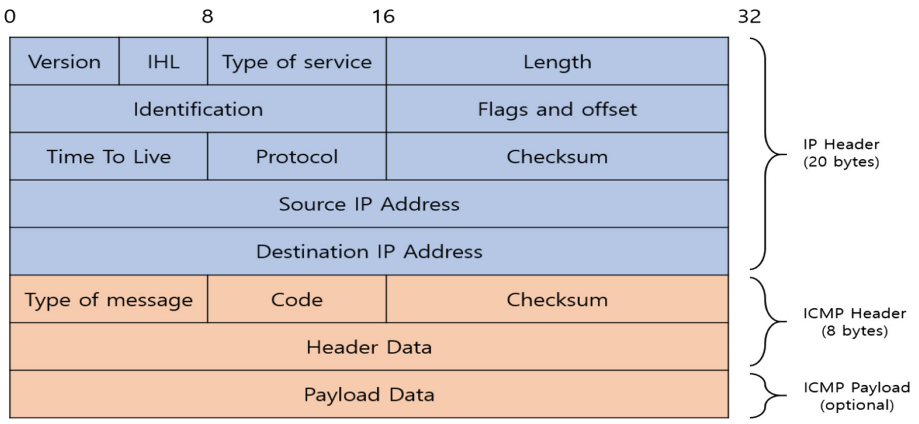


Fig. 1. ICMP message format

The ICMP consists of IP Header, ICMP Header and ICMP Payload, as shown in Fig. 1. The components of IP Header and ICMP Header are defined in Tables 1 and 2, respectively.

Table 1. IP header

Filed	Explanation
Version	IP version
Header length	IP header length
Type of service	IP define packet priority
Total length	Overall length including header and data length
Identification	Packet identifier
Flags	Fragmentation bit
Fragmentation offset	Order portion of the fragment
Time to live	Packet lifetime
Protocol	Protocol
Header checksum	Header check
Source IP address	Source IP address
Destination IP address	Destination IP address

Table 2. ICMP Header

Filed	Explanation
Type	Message type
Code	Message information
Checksum	ICMP header check

The ICMP includes the information presented in Tables 1 and 2 and can send data to the ICMP Payload. However, the Payload is not used in the case of ICMP message types [2]. The analysis is carried out using five operating systems, i.e., Windows (7, 8.1, 10), Ubuntu (Linux 18.x, Debian 7.x), and Wireshark [3].

4 Analysis

The network packet is extracted using Wireshark, and the items are as shown in Table 3. The information presented in Table 3 can be extracted from IP Header and ICMP Header. Next, using Wireshark, we check the packet information of Windows and Ubuntu. As shown in Fig. 2, data collection is performed by using ICMP communication, using the Windows Ping command on Windows. The collected information is organized based on the information received by the ICMP response.

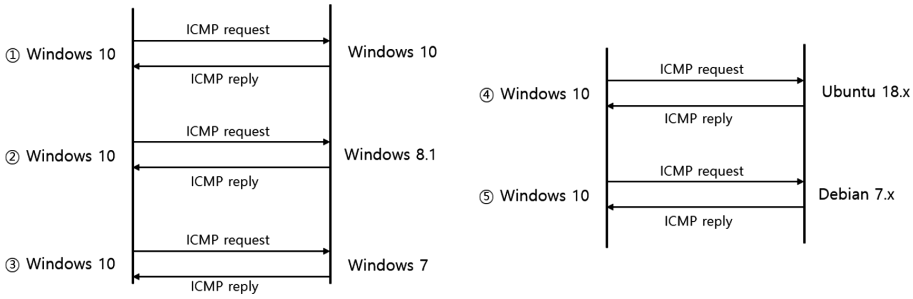


Fig. 2. Order of data collection

Table 3. IP header and ICMP extraction items

Version	Header length
Type of service	Total length
Identification	Flags & Frag offset
Time to live	Protocol
Header checksum	Source IP address
Destination IP address	ICMP Type
ICMP code	ICMP checksum

Table 4. Data collection information

	①	②	③	④	⑤
Version	4				
Header length	20				
Type of service	0x00				
Total length	60				
Identification	0x1749	0x187b	0x6fa4	0xb75e	0x3f3e
Flags & Frag offset	0x0000				
Time to live	128			64	
Protocol	ICMP				
Header checksum	0x5eXX	0x8cXX	0x1bXX	0x6eXX	0xeXXX
Source IP address	168.188.x.x				
Des IP address	168.188.x.x				
ICMP type	0 (reply)				
ICMP code	0				
ICMP checksum	0x19ce	0x19c2	0x19be	0x554f	0x5527

In Table 4, ①, ②, and ③ correspond to the Windows series, and ④ and ⑤ correspond to the Ubuntu series. From the IP Header, it was confirmed that the version, header length, service, and total length are the same for all Oss; however, the

identification, time to live (TTL), and header checksum are different. Among them, time to live (TTL) is the parameter that can classify the operating system, and we can confirm that Windows and Ubuntu. If the TTL is 128, the OS is determined to be Windows, and if the TTL is 64, the OS is Ubuntu. After classifying the two operating systems with TTL, we attempt to determine the versions for each current operating system.

There are three parameters that are different for each OS: Identification, Header Checksum, and ICMP Checksum. Checksum cannot be used because it can be obtained by adding bytes of other items besides Checksum and taking 1's complement. Therefore, only the Identification is possible except for the TTL.

4.1 Identification Analysis

Identification serves to uniquely identify each IP packet as the sole identifier of the packet. Fragmentation can be configured to disable this fragmentation by using the Flags bit because the length of the transport packet is different for each physical medium. If fragmentation fails, this packet is not used. We attempted to determine whether the Identification parameter can provide results for each operating system. The method involved collecting ICMP messages for each operating system four times. The results are presented in Table 5.

Table 5. Identification collection information

	Win 10	Win 8.1	Win 7	Ubuntu 18.x	Debian 7.x
1 회	0x6432	0x59c2	0x0083	0xccc4	0x4f08
2 회	0x6434	0x59c3	0x0084	0xcce7	0x4fed
3 회	0x6436	0x59c4	0x0085	0xcd1e	0x4ff1
4 회	0x6438	0x59c5	0x0086	0xcd27	0x50ab

Show in Table 5, we can see that Identification increased by 2 for Windows 10 and by 1 for Windows 8.1 and Windows 7. And Ubuntu can be seen to be irregular. Table 6 summarizes the above information.

Table 6. Results

TTL 128	Identification 2↑	Windows 10
	Identification 1↑	Windows 8.1, Windows 7
TTL 64	Irregular	Ubuntu 18.x, Debian 7.x

5 Conclusion

In this study, we examined whether it is possible to distinguish the OS by using ICMP. We were able to classify operating systems using TTL and Identification fields. However, as in Windows 8.1 and 7, the specific values could not uniformly increase. Further, Linux systems had irregular occurrence of Identification fields, and thus, classification was not possible. However, since these phenomenon are problems caused only by ICMP, it is necessary to advance research using TCP rather than ICMP. In addition, five operating systems have been set up, but it is necessary to study what types of operating systems such as Unix, BSD, and Windows XP are classified.

Acknowledgments. This research was supported by the Ministry of Science and ICT (MSIT), Korea, under the Information Technology Research Center (ITRC) support program (IITP-2018-2016-0-00304) supervised by the Institute for Information & Communications Technology Promotion (IITP).

References

1. Humer, S., Murphy, A.: OS fingerprinting techniques and tools, cryptography and network security. Keene State College, CS-455 (2013)
2. Bellovin, S.M., Leech, M., Taylor, T: ICMP traceback messages (2003)
3. Lamping, U., Sharpe, R., Warnicke, E.: Wireshark User's Guide for Wireshark 2.1 (2014)

A Study of a Common Network Behavior Detection System Using Remote Live Forensics

Jeonghoon Seo and Yoojae Won(✉)

Department of Computer Science, Chung-Nam University,
Daejeon, South Korea
{sjh9309,yjwon}@cnu.ac.kr

Abstract. Recently, attackers spread malware mainly through networks to infect more hosts faster. However, it is difficult to track an attacker as network records are highly volatile. Therefore, there is an increasingly growing interest in live forensics which collects and analyzes evidence from the host system in an active state. This paper proposes a system that detects common behaviors based on network access records from multiple endpoints using a remote live forensic tool.

Keywords: GRR · Network malware · Live forensics · Behavior-based detection

1 Introduction

Live forensics is a technology that collects and analyzes the necessary digital evidence while maintaining the integrity of the evidence from the active target system [1]. Attackers have recently developed improved attacking techniques that do not leave their traces on volatile data in order to evade tracking by investigators. Traces of such attacks exist in volatile data, including RAM; thus, interest in live forensic techniques is increasing.

People first started developing malware simply for fun or to demonstrate their ability. However, there are increasing incidents of attacks targeting multiple PCs in companies or organizations for monetary gain or interest [2]. Also companies and other organizations use antivirus programs to reduce damages. However, the development of mutant malware that bypasses them is also on the rise.

Malware is frequently spread through network activities, including visiting web-pages, emailing spam, and downloading files. In many cases, after the target PC is infected, it is connected to the attacker's C&C server to be commanded or controlled by the attacker [3]. Investigating network activity records of infected PCs may reveal the common propagation path of the malware or common communication records with the attacker's C&C server. If an attack using this kind of malware is conducted in a company environment with multiple PCs connected to the same network, the source of malware propagation in the infected PCs or common network communication with the attacker's C&C server can be detected. This may be a critical clue to detect attacks and respond quickly.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 291–296, 2020.
https://doi.org/10.1007/978-981-13-9341-9_50

This paper proposes a system to detect attacks targeting multiple PCs; the system collects network connection records of PCs and detects common network behaviors between them using a remote live forensic tool GRR.

2 GRR

GRR is an abbreviation standing for GRR Rapid Response. It is a remote live forensic tool for responding to security incidents. GRR, aiming at fast and expandable remote survey and analysis of the host, was developed to conduct forensic surveys on numerous systems simultaneously [4].

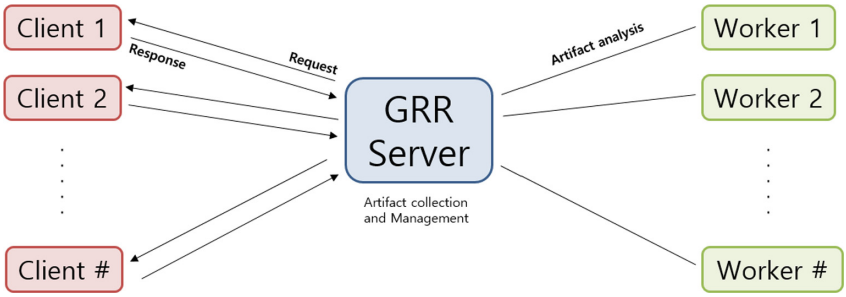


Fig. 1. GRR scheme.

Figure 1 shows the scheme of GRR. GRR consists of client, server, and worker. GRR is installed on a client as an agent, and it exchanges messages with the server through the HTTP protocol after installation. The server receives responses from each client and stores the data in a database. A worker plays the role of an analyzer, gives a command to collect a client's artifacts through the server, and analyzes the collected artifacts. Multiple clients can be connected to one server, and multiple workers can be simultaneously connected to the server to perform analysis [5].

A flow is a process in which the server collects artifacts from a client PC upon a worker's command. When the worker remotely selects and executes a desired flow, the request is sent to the client from the server [6]. The client who receives the request extracts the requested artifacts and sends them to the server, and the worker can access and analyze the artifacts in question anytime anywhere.

A hunt refers to the sending of a flow, which is the function that extracts and collects artifacts from one client, to all or some systems managed by the server. In other words, artifacts can be simultaneously collected from multiple clients through a hunt. When a hunt is executed, a client rule is written, and the request is sent exclusively to clients that match the rule [7].

3 Common Behavior Detection System Using Remote Live Forensics

This paper proposes the following common behavior detection system using remote live forensics, aiming to detect attacks conducted through a network on multiple hosts, such as companies or other organizations.

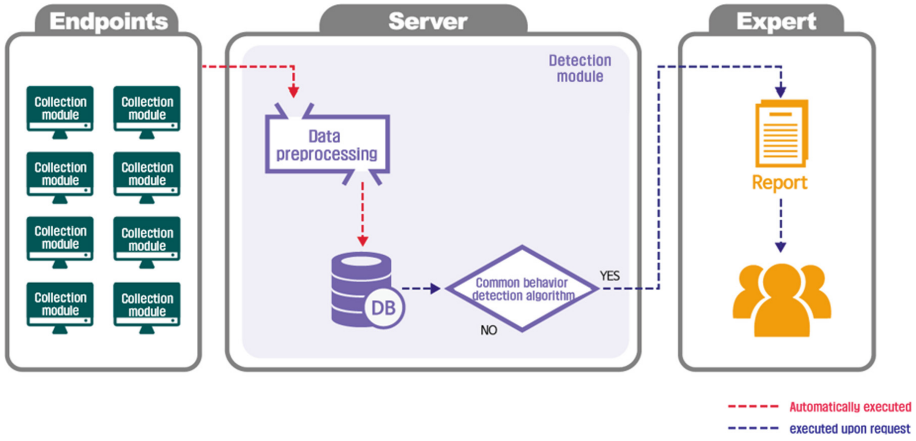


Fig. 2. Scheme and communication of the common behavior detection system.

Figure 2 shows the structure and communication process of the proposed system. The system consists of a collection module that collects network connection records, a data preprocessor, and a detection module that operates a database and detects common network behavior through an algorithm. The results of the detection are reported to the experts and assist the expert's judgment.

3.1 Collection Module

The collection module extracts the network access records of each client using the GRR hunt function and sends them to the server. The collection module automatically execute Netstat Hunt using GRR API through code. As a result of collection, a JSON File containing network access records of each client is collected, and it is sent to the detection module.

Although additional network equipment and tools could be used to collect network access records, GRR was used for collection modules because the Server-Client structure allowed the server to manage multiple endpoints and easily collect data and retain data for long periods of time.

3.2 Detection Module

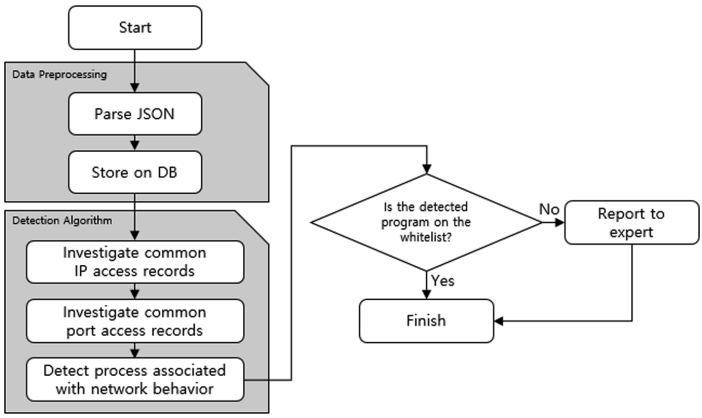


Fig. 3. Flowchart of the detection module.

Figure 3 presents a flowchart of the detection module. The detection module consists of a data preprocessor, a database, and a common behavior detection algorithm. The data preprocessor sorts the data needed for common behavior detection from the JSON File received from the collection module. There are about 40 attributes in the JSON File collected by the collection module. Among them, Client Name, Timestamp, Source IP and Port, Destination IP and Port, and PID are helpful for detecting common actions. Therefore, it is necessary to process other attributes that are not helpful for common action detection. The data preprocessor parses the JSON File to delete useless attributes and to extract only needed attributes. The extracted data are saved and stored on the database established in the server.

The common behavior detection algorithm uses the data stored in the database to check whether the clients record the connection to the same IP and port, the existence of a common process, and the network communication. Through this process, it is possible to detect a common network connection behavior among a plurality of clients.

However, this detection algorithm shows a high error rate because the range to cover is too vast. A typical example is the system process. The problem is that the system process is always detected since it exists for all clients. Our proposed system applies the whitelist technique to address this problem. To apply the whitelist, it requires an assumption that sufficient data have been accumulated through data collection starting from a clean state when the clients had not been infected by malware. Under such an assumption, verified safe processes are added to the whitelist and will not appear in future detection results.

The detection results of the detection module consist of IPs and ports that are connected to the process as well as the number of commonly detected clients. All detection results are printed and reported to an expert. The expert determines whether to block or allow the processes in question based on the results.

4 Common Network Behavior Detection Test

4.1 Test Environment and Scenario

A test was conducted to confirm that common network behaviors are detected through malicious code communicating with the attacker's C&C server and network. The test environment was set up as follows. Collection module was installed in three virtual machines and a detection module was installed in one real machine. It also used one real machine to act as the attacker's C&C server and malicious code called Babylon RAT. Babylon RAT is configured to communicate with the attacker's C&C server on port 9999.

In the test scenario, first, the Babylon RAT is executed in three virtual machines, and the collection module is operated to collect and store the volatile network connection record and the running process list. It then checks to see if the communication records between the Babylon RAT and the attacker's C&C server are detected as common network behavior by the common module detection algorithm of the detection module.

4.2 Experiment Results

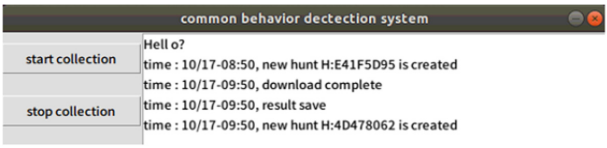


Fig. 4. Screenshot of the collection module in operation.

Figure 4 shows the operation of the acquisition module. After the collection is completed, the results are downloaded and stored, and a new collection is started. The collection period is 60 min and repeats until you click Stop collection.

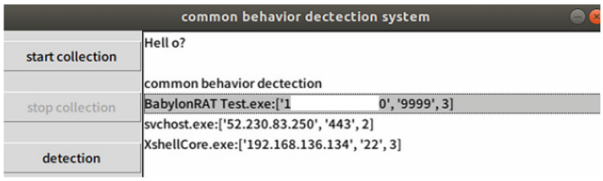


Fig. 5. Screenshot of detection results.

Figure 5 shows the results of common behavior detection based on collected data. The Babylon RAT malware Babylon RAT Test commonly communicated with the attackers' C&C server on the 3 clients through port 9999.

5 Conclusion

This paper proposed a common behavior detection system using network access records to detect attacks conducted through networks. It was confirmed that the system detected a common network behavior in clients via a test using malware.

The proposed system collects network access records from multiple clients, detects common points in them, and reports the results to an expert. This system is expected to rapidly detect attacks that infect multiple hosts in companies and other organizations, communicate with the attacker's C&C server, and execute additional commands or steal internal information.

In the future, additional studies will be conducted to detect common points in behaviors occurring in systems as well as attacks launched through networks. Finally, we plan to go beyond simple detection; we will conduct mobile and static analysis through the process or active memory dump and provide experts with results to help them determine the existence of malware.

Acknowledgments. This research was supported by the Ministry of Science and ICT (MSIT), Korea, under the Information Technology Research Center (ITRC) support program (IITP-2018-2016-0-00304) supervised by the Institute for Information & Communications Technology Promotion (IITP).

References

1. Lim, K.S., Park, J.H., Lee, S.: Trends and challenges of current digital forensics. J. Sec. Eng. **5**(6), 461–474 (2008)
2. Egele, M., Scholte, T., Kirda, E., Kruegel, C.: A survey on automated dynamic malware-analysis techniques and tools. ACM Comput. Surv. **44**(2), 6 (2012)
3. Son, K., Lee, T., Won, D.: Design for Zombie PCs and APT attack detection based on traffic analysis. J. Korean Inst. Inf. Secur. Cryptol. **24**(3), 491–498 (2014)
4. GRR rapid response documents. https://grr-doc.readthedocs.io/en/latest/
5. Cruz, F., Moser, A., Cohen, M.: A scalable file based data store for forensic analysis. Digit. Investig. **12**, 90–101 (2015)
6. Cohen, M.I., Bilby, D., Caronni, G.: Distributed forensics and incident response in the enterprise. Digit. Investig. **8**, 101–110 (2011)
7. Moser, A., Cohen, M.I.: Hunting in the enterprise: forensic triage and incident response. Digit. Investig. **10**, 89–98 (2013)

A Rationalization of Hangul Codes Based on Jeongeum Principle in ISO/IEC10646BMP

Jeongyong Byun(✉)

Computer Engineering Department, Dongguk University, DongdaeRo 123, Gyeongjusi, Gyeongbuk, Korea
byunjy@dongguk.ac.kr

Abstract. We find and propose a strategy to rationalize three kinds of Hangul code existing in ISO/IEC 10646BMP by assessing the syllable expression ability and also applying four rules of explanation of combining letters according to the scientific principle of Hunminjeongeum.

Keywords: ISO/IEC 10646BMP · Unicode · Optimization · Rationalization · Jamo · Jaso · Syllable · Combining · Attaching · Compositing

1 Introduction

ISO/IEC 10646-1:1993 UCS (four octets) [1] becomes international standard in 1993 based on the second edition of DIS 10646. Its basic multilingual Plane (BMP) is equivalent to Unicode with only one plane. There are three Hangul codes [9] under the code charts of the East Asian scripts [9] such as Hangul Jamo (U+1100), Hangul compatibility Jamo (U+3130) and Hangul syllable (U+AC00).

International standards can adopt only national standard. KS C 5601:1974 stands for Hangul compatibility Jamo and KS C 5601:1987 or KS X 1001:2004 [10] support completion syllable code. Hangul Jamo actually means grapheme of syllable character and another naming is *Jeongeum* or *Jaso* [7]. It became a national position in preparation for the ISO/IEC TG97 SC2 Seoul meeting in 1992 and was adopted as draft national standard in March, and finally adopted as one code of UCS BMP at the meeting in June. It is named *Hangul Jamo*. Three kinds of Hangul codes exist in BMP for Korean writing system. KS X 2016-1:2007 gives a guidance but it may be mixed.

In this paper, we evaluate that three kinds of Hangul codes in BMP need to be maintained according to the scientific principle of Hunminjeongeum (Jeongeum) [7]. Also, we consider the problems that can occur when using the three codes together. We propose a strategy to rationalize Hangul codes in BMP by assessing which Hangul code is suitable for all applications.

2 Previous Works

There are some studies [2] to optimize Hangul code using scientific principle of Hunminjeongeum (Jeongeum). The Jeongeum code [7] is derived in 1991 and offered to domestic JTC1/SC2/WG2 in January 1992, and is decided in national standard in

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 297–302, 2020.
https://doi.org/10.1007/978-981-13-9341-9_51

March. It was adopted as an international standard at the Seoul meeting in June and is also included in the BMP. There was a study [5] to utilize Hangul Jamo code as a writing system of Hangul and Jeongeum notation. A promising study [4, 6] is to develop IME which expresses 39.9 billion syllable fonts by applying 4 rules with 45 characters. And research [3] raised the problem of the current Hangul Jamo code.

3 Hangul Coding Scheme

Hangul is phonetic, syllabic, and a phoneme letter. If we understand the scientific principles of Jeongeum well, we can know that they can express the smaller set of basic letters the more set of syllable.

3.1 Scientific Principle of Jeongeum

Jeongeum is originally defined as forty-five basic letters and four rules. A syllable of Jeongeum consists of the initial, the middle, and the final letter. The number of the initial sound is 17, the middle is 11, and the final is equal to the initial. Four rules are defined in the explanation of letter combination such as light sound rule, combining rule, attaching rule, and composing rule. The combining rule extends the member of three components by combining two or three letters side by side and the results are 5219 in the initial (*In*), 1463 in the middle (*Mn*), and 5220 in the final (*Fn*), where n stands for the number of combined letters.

$$In = \sum_{k=1}^{3} x^k,\ Mn = \sum_{k=1}^{3} y^k,\ Fn = \sum_{k=1}^{3} z^k \quad \text{for } x = 17,\ y = 11,\ z = 17 \tag{1}$$

$$Hn = In * Mn * Fn \tag{2}$$

The composing rule stands for that the initial, middle, and final letters should be composed to form syllables. After evaluating Eq. (1), *Hn* in Eq. (2) composes 39,856,772,340 (about 39.9 billion) syllables. It is the working space of Jeongeum and an astronomical number. This can be a proof of which Jeong Inji's preface says that if there are sounds natural to Heaven and Earth there must be letters natural to Heaven and Earth. The light sound rule and attaching rule are related to positioning ung sound ㅇ and the vertical or horizontal middle letters.

3.2 Jamo Code

Hangul Jamo stands for consonant and vowel letters defined by Hangul spelling, which calls the final letter saucer consonant. This is the first Hangul code, KS C 5601:1974, which grants codes for 30 consonants and 21 vowels. The number of the consonant is divided into the initial consonant 19 and the final consonant 28 including empty. If applying these letters for Eq. (2) the size of syllables produced from Jamo code becomes 11,172 syllables.

3.3 Syllable Code

Hangul can take two coding schemes when coding syllables. One is the combination syllable code that was revised in 1982 as national standard. MSB of 16 bits is a sign for Korean or English, and the remaining 15 bits are separated into three by 5 bits, of which values are 32. It can contain 19 consonants, 21 vowels, and 28 consonants including empty. However, this violates the ISO/IEC 2022 code extension.

Another one is the combination syllable code, KS C 5601:1987 or KS X 1001:2004, that was adopted in 1987 as Korea standard. According to ISO/IEC 2022, 94 characters are placed within graphic right area (GR). If using 2 bytes for Hangul, the size becomes 8836 (94 * 94) and it is smaller than 11172 syllables. By the way, Hangul 2350 syllables are assigned to GR 0xB0A1 - 0xC8Fe and puts 94 user-defined sections. 2350 Hangul syllables are only 21% of 11172 syllables. It lacks 350 syllables for Korean language life and can be supplemented from 1930 in KS X 1002:2001. It results in Hn = 4280 (38.3% of 11172) but it is still incomplete.

3.4 Unified Hangul Code

Even after the representation of the syllabic code in 1987, Word processor Hancom Haangul has been kept to use a combining syllable code system instead of the standard code system. Table 1 shows the unified Hangul code.

Table 1. Assigned areas of 8822 in Unified Hangul Code

	1st byte	2nd byte			Subtotal	Total
Zone 1	81-A0	41-5A	61-7A	81-FE		
	32	26	26	126	5,696	
Zone 2	A1-C5	41-5A	61-7A	81-A0		
	37	26	26	32	3,108	
Zone 3	C6	41-52				
	1	18			18	8,822

MS Word, on the other hand, follows the standard but most users avoid to use it. To solve this problem, Microsoft assigned 8,222 syllables to some areas using CP949 (MS949). It is called Unified Hangul Code [10]. However, the illegal assignment for 8822 is not in compliance with ISO/IEC 2022 [8] and not registered on IANA. The violations are as follows. First, 0x81–0x9F of 0x81–0xA0 in the first byte cannot be used because it is the control character area of GR. Second, in the second byte, it invaded the Roman uppercase zone, 0x41–0x5A and Roman lowercase zone 0x61–0x7A of GL, and the control character zone 0x81–0x9F of GL. It stays at Hn = 2350.

4 Code Assessment

Hangul generates syllables by the combining rule and composing rule with basic letters. One of the important factors in evaluating each code is to apply the assessment of syllable generation ability.

Table 2. Syllable size of Hangul Codes

Type	Byte	Comments	Syllable
Jamo	1	Consonant 30, Vowel 21	11,172
Combined syllable	2	Consonant 19, Vowel 21, Final 28	11,172
Completed syllable	2	No grapheme, KS X 1001	2,350
	2	No grapheme, KS X 1002	1,930

Three kinds of Hangul code exist in the ISO/IEC 10646BMP. First, completion syllable code is KS X 1001:2004 (2350) and KS X 1002:2001(1930) and their total is 4280 (38.3% of 11172), but BMP assigned 11172 as shown in Table 2. Second, Hangul Jamo refers to Jeongeum code. Half width Hangul Jamo is unknown on useful usage. Hangul Compatibility Jamo consists of 30 consonants and 21 vowels, but separating consonants into the initial and the final and applying Eq. (2) can generate 11,172 syllables. Combining rules in vowels and consonants were limited to double. Compatibility Hangul Jamo in the BMP encodes consonants and vowel sets, which correspond to a subset of Hangul Jamo. The Jamo type refers to the coding of the initial, middle, and final letter of a syllable (Table 3).

Table 3. Jaso-type codes in BMP

	1100	1100	A97F	D7FF	Fill	Total
Initial	90	5	29	0	1	125
Middle	66	5	0	23	1	95
Final	82	6	0	49	0	137
Sum	238	16	29	72	2	357

Jaso code originally refers to Jeongeum code. It can compose Hn = about 39.9 billion syllables by applying the initial, the middle, and the final to the four rules including the combining rule according to Jeongeum. This is an astronomical number.

In BMP, no combining rule is used at all. So Hangul Jamo has a double or triple letters set for the initial, the middle, and the final. The listing order is to place the modern Hangul Jaso first and then the old Hangul. The total number of Jaso assigned to the extension-A and B is 357 characters. If applying Jaso set such as {In = 125, Mn = 95, Fn = 137} to Eq. (2), Hn becomes about 1.63 million syllables.

5 Rationalization Strategy

Three kinds of Hangul codes existing in BMP are arranged in 6 areas. A good strategy to rationalize them is that it must follow the scientific principle of Jeongeum and cover the complete set of syllables by the principle. If we recover four rules of Jeongeum, we select only single letter 45 as basic letters and remove 312 double or triple letters. If applying 45 letters for four rules, 39.9 billion syllables can be composed completely. It shows that Jeongeum can generate the largest syllable set among three Hangul codes. The number will Then we can return 4 blocks such as Hangul Compatibility Jamo (U+3130), Hangul syllable (U+AC00), Hangul Jamo extension-A, Hangul Jamo extension except Hangul Jamo (U+1100) area.

6 Conclusion

Hangul Jamo has the highest ability that can apply the scientific principle of Jeongeum if code uses only 45 single letters and four rules. Then it can be the best code to rationalize three kinds of Hangul codes assigned to BMP. And also KS X 2016-1 is not required. The four rules say that Jeongeum must use software. It means that Jeongeum or Hangul is the most relevant letters to computer system in the world. So it is desirable to return 5 blocks such as U+3130, U+AC00, extension-A and B.

In the further research, we are going to identify and develop common functions of various algorithms and text editors required in processing Jeongeum Jaso 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 (NRF-2015R1D1A1A01060408).

References

1. ISO/IEC 10646 UCS. http://www.w3c.or.kr/i18n/hangul-i18n/iso10646.html
2. Byun, J.Y., Oh, S.H.: Advances in multimedia modeling: a representation of Korean syllabic characters defined in Hunminjeongeum. In: Cham, T.J., Cai, J., Dorai, C., Rajan, D., Chua, T.S., Chia, L.T. (eds.) MMM 2007. LNCS, vol. 4351, pp. 662–670 (2007)
3. Kim, C.H.: A problem of KS completion Hangul code. Korean Life, Autumn, Korea (1989)
4. Byun, J.Y., Hong, S.B., Kim, H.Y.: Advances in computer science and ubiquitous computing: a design scheme of combined syllable fonts for Hunminjeongeum. In: Park, J.J. (J.H.), Pan, Y., Yi, G., Loia, V. (eds.) CSA-CUTE 2016. LNEE, Bangkok, Thailand, vol. 421, pp. 1096–1102 (2016)
5. Kim, G.Y., Byun, J.Y.: A glyph reduction evaluation of combined Jeongeum fonts. In: PDCAT 2018. Korea Information Process Society, Jeju (2018)
6. Byun, J.Y.: Hunminjeongeum principle and Hangul Code. KIISE/SigLE Rev. **1**(2), 3–7 (1992)

7. Lee, H.J., Byun, J.Y.: A web input method for full set syllables defined by Hunmin-jeongeum. J. KIISE: Comput. Pract. Lett. **19**(6), 371–375 (2013)
8. ISO/IEC 2022. https://ko.wikipedia.org/wiki/ISO/IEC_2022
9. Code Charts, Unicode. http://unicode.org/charts/
10. KS X 1001:2004. https://ko.wikipedia.org/wiki/KS_X_1001

Design Automation System for Review Analysis Affiliation for Online Educator Reliability Prediction

Kihoon Lee[1], Hyogun Kym[2], and Nammee Moon[1(✉)]

[1] Department of Computer Engineering, Hoseo University, Asan, South Korea
lkh51738@gmail.com, nammee.moon@gmail.com
[2] Center for Business Art, Ewha Womans University, Seoul, South Korea
Kym@ewha.ac.kr

Abstract. In this paper, we designed a review analysis automation system to grasp the credibility of the online education matching platform. Web crawling collects and parses reviews and ratings of educators who are atypical data. We will build an emotional dictionary based on the educational field to grasp online educator credibility using collected review data and SO-PMI. We also propose a method for building large - scale learning data based on the emotion dictionary constructed. We proposed a system that provides more reliable review analysis results by measuring the accuracy of emotion dictionary by using deep learning in constructed learning data and evaluation data. Through this, we intend to help judge the credibility of educator in O2O education matching.

Keywords: Deep learning · RNN · LSTM · SO-PMI

1 Introduction

The 21st century is called the knowledge and information society, and many strategists and scholars emphasize their roles in the knowledge and information society and claim that human resource discovery and management is an important task in the 21st century [1]. In the field of human resource development, past collective education and technology-oriented education are gradually diminishing. On the other hand, education in one-to-one relationships tends to increase. In other words, education that considers both personality and knowledge is attracting attention [2].

In the past, if you were able to match this one-to-one relationship with offline advertising or a community-style platform, there are now a number of O2O educational matching platforms in the spotlight. There is a study that offline education is more effective in education of specific fields such as arts and physical education than online. Due to these characteristics, online matching is frequently performed offline. When such offline education is conducted, the reliability of the tutor and the tutu can be a very important matching factor [3].

In this paper, we will use the review analysis of educators to understand the reliability of online educators. Recently, review analysis has been used in various fields, and work on numerical evaluation of reviews is actively proceeding [4, 5].

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 303–309, 2020.
https://doi.org/10.1007/978-981-13-9341-9_52

Among them, the evaluation of the rating using the deep-learning is the most active. Deep-learning is a technique of learning and constructing models using train data and classifying test data through such models. The deep-learning technique can build a good model to learn with a sufficient amount of learning data. However, it is not easy to obtain good enough amount of learning data to build a good model.

In this paper, we propose a method of constructing emotional dictionaries by educational field using SO-PMI to solve these problems. Using the built-in emotional dictionary, the learning data can be extracted for the non-digitized review. Also, we propose a model that learns the collected learning data using LSTM, which is a deep-learning model, and then evaluates the accuracy and results by applying evaluation data that can be used to extract scores.

2 Related Research

2.1 SO-PMI

SO-PMI is a technique that reduces the bias of individual words in PMI. PMI is abbreviation of Point-wise Mutual Information. It is a technique to quantify the correlation between two points x and y by applying probability theory. PMI is based on Mutual Information. $PMI(x, y)$ between two points for two points x, y with each occurrence probability $P(x)$, $P(y)$ can be defined as follows.

$$PMI(x, y) = log \frac{P(x, y)}{P(x)P(y)} \tag{1}$$

The probability of simultaneous occurrence of x and y assuming that x and y occur independently is compared with the probability of simultaneous occurrence of measured x and y to determine how much the two variables have correlation. The equation for calculating SO-PMI is shown in Eq. 2 below.

$$SO - PMI(x) = \sum_{px \in PX} PMI(x, px) - \sum_{nx \in NX} PMI(x, nx) \tag{2}$$

Where PX is a set of positive reference words, and NX is a set of negative reference words. $SO - PMI(x)$ means the result obtained by subtracting the PMI sum of the negative word set from the PMI sum of the word x and the positive word set. In this paper, we use the SO-PMI technique to construct emotional dictionaries for each educational field.

2.2 Long Short Term Memory (LSTM)

The LSTM is a model that compensates for the degradation of the learning ability due to the reduction of the gradient value at the back propagation as the length occurring in the RNN (Recurrent Neural Network) becomes longer. Also, it is a deep- learning model that shows good performance in sequential data and can learn bi-directionally on input data. In this paper, we will use the SO-PMI emotional dictionaries to learn the

rated reviews as learning data and then test the accuracy by applying the test data with the rating from the existing online education matching platform.

3 The Proposed System

3.1 Online Educator Reliability Forecasting System

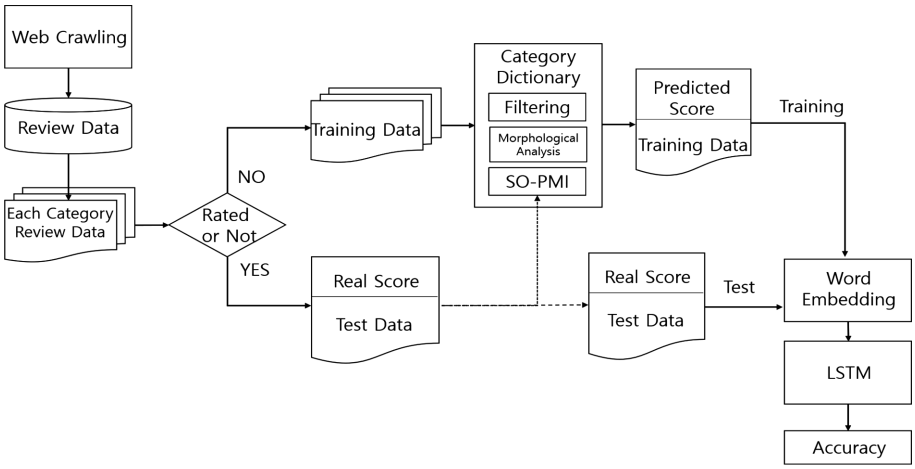


Fig. 1. Total system processing

The proposed learning data construction process is shown in Fig. 1. We use web crawlers to collect reviews of educators across educational platforms on a platform that provides online training matching services. At this time, if the rating service is provided for the review, extract it with the review. The extracted review is divided into learning data and evaluation data according to the presence or absence of a rating. After dividing the review data that does not have a rating into the filtering process and the morpheme unit, the morpheme of the top 30% rating of the test data and the morpheme of the lower 30% rated sentence are taken as the reference word set, and the emotional dictionary. You can assign a rating to the sentence of the training data using the built-in emotion dictionary. Then, the LSTM emotion analyzer is trained with the training data given the rated points.

3.2 Data Preprocessing

In this paper, we collected review data on educators by education field on existing online education matching platform to propose a system to identify reliability through review of online educators. The following Table 1 is an example of a pre-process for the crawl data.

Table 1. Examples of changes in input data before and after preprocessing

Text	Rating
[“아직”, “노래”, “무섭지만”, “차근”, “나아지”, “는”, “모습”, “에”, “재밋”, “게”, “수업”, “진행”]	4
[243, 14583, 2, 12924, 54580, 2382, 299, 3855, 800, 92815, 74, 243]	4

3.3 Word Polarity Judgment

In order to construct emotional dictionary, only ‘adjectives’ were used in this study. Prior to the extraction of adjectives, we removed unnecessary words and meaningless words from the web crawling process and removed the abbreviations such as numbers and special symbols. Finally, the SO-PMI scores of the extracted adjectives are calculated to determine the polarity of the words. PMI is a score indicating the similarity between two words. If the PMI value is high, the similarity between the two is high. The calculated PMI value is used for SO-PMI calculation. Prior to the application of SO-PMI, the set of positive and negative reference words in the analysis data is set, and the SO-PMI score, which is the score obtained by subtracting the PMI sum from the negative reference word and the sum of the PMI and the positive reference word, is calculated. If the score of SO-PMI is high, it is similar to the affirmative reference word set. On the contrary, it can be interpreted as positive because the affinity with the negative reference word set is low. Likewise, if the SO-PMI score is low, it can be interpreted as a negative word. Since the SO-PMI value is influenced by the set of reference words, it is important to set the reference word set appropriately. In this paper, we try to set the upper 30% as positive word set and the lower 30% as negative word set by using the average of test scores and the number of reviews to make reference word set. The frequency of the words generated from the two sets is compared and the frequency of the word is attributed to the higher set. The generated positive and negative top 15 words were used as reference word groups. The number of reference word groups was based on previous studies [6].

3.4 Emotional Dictionary Construction and Accuracy Determination

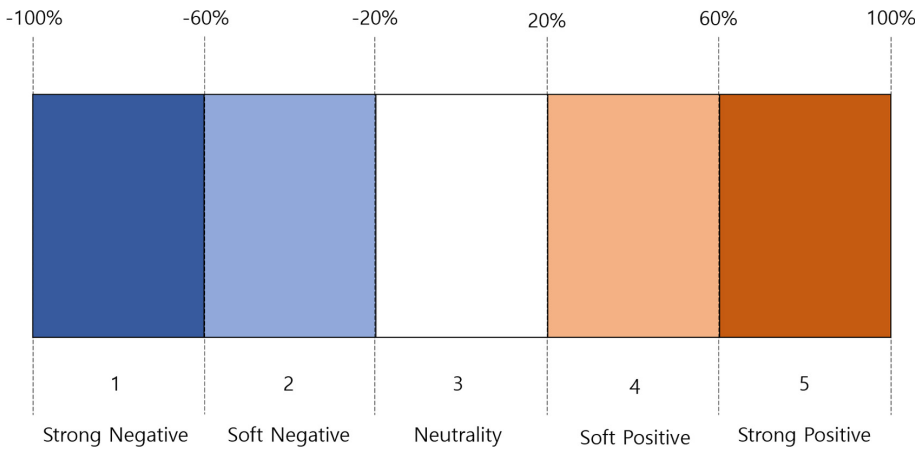


Fig. 2. Classify SO-PMI score

After determining the word-by-word polarity score through SO-PMI, emotional dictionary is constructed based on this. In the emotional dictionary, each word is classified into 5 strong positive, weak positive, strong negative, weak negative, and average 5 levels of positive negative. If SO-PMI is greater than 0 and the value is greater than or equal to 0, it is defined as 'strong positive'. If the SO-PMI is greater than 0, 'strong positive'. The rules for classifying word polarity based on SO-PMI are shown in Fig. 2.

Table 2. Adaptive sentiment dictionary

Rank	Text	Sentiment Score
1	노래	4.14821
2	재미	3.52172
3	좋다	4.51275
4	매우	3.65174

Table 2 seeks to make more accurate reviews and predictions through the establishment of emotional dictionaries by field as a result of the establishment of emotional dictionaries by field.

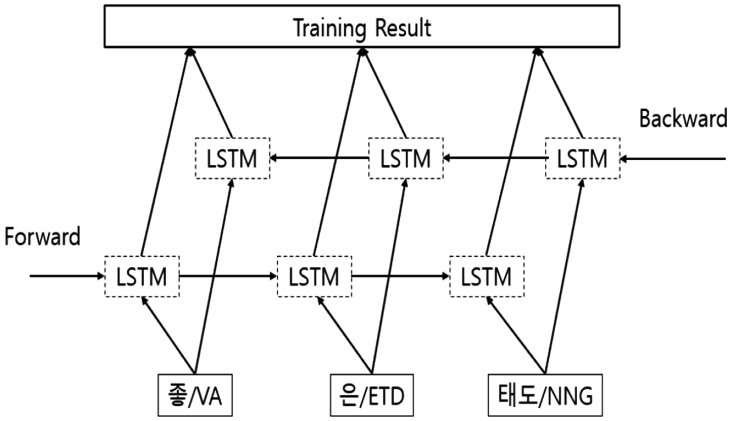


Fig. 3. LSTM example

Figure 3 is an example of the LSTM model proposed in this paper. The result of word embedding learning data is learned as an input value, and the learned result is applied to the evaluation data to confirm the accuracy and the result. Accuracy measurements are made through a comparison of the predicted point value and the actual point value from the LSTM learned in advance for the review of the evaluation data.

4 Conclusion

In this paper, we have designed an automated system to grasp the credibility of the tutor when using the online education matching platform. We used the SO-PMI technique to define the polarity of the tutor reviews, used it to build the emotion dictionary. We proposed a system that helps to collect training data used in LSTM emotional analysis by predicting and assigning ratings to reviews that do not have ratings by using the emotion dictionaries for each educational field. However, the research limitations of this study are as follows for the customized emotional dictionary construction required for the emotional analysis of each field.

In this paper, we tried to use only the adjectives that are expected to have a pronounced polarity in the review of the collected emotional dictionaries using training data. However, it is expected that words of other parts such as verbs and adverbs can be used for emotional analysis, and research on emotional dictionary construction including various parts of speech should be carried out.

Acknowledgement. This research is supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program 2018 (R2018020083).

References

1. Choi, Y.J., Lee, S.J., Jeong, J.: Substitute and complementary relationships between online and offline courses in foreign language education. Korean J. Econ. **25**(1), 45–60 (2018)
2. Kim, S., Lim, K.Y.: The moderating effects of perceived usefulness and self-regulated learning skills on the relationship between participative motivation and learning satisfaction in online continuing education programs. J. Lifelong Learn. Soc. **13**(3), 85–107 (2017)
3. Zhang, P., Moon, H.C.: A study on the effects of O2O commerce characteristics and consumer characteristics on trust, desire and behavioral intention in China. In: Korea Trade Research Association Conference, pp. 107–123 (2015)
4. Kennedy, A., Inkpen, D.: Sentiment classification of movie reviews using contextual valence shifters. Comput. Intell. **22**(2), 110–125 (2006)
5. Lee, S.H., Cui, J., Kim, J.W.: Sentiment analysis on movie review through building modified sentiment dictionary by movie genre. J. Intell. Inf. Syst. **22**(2), 97–113 (2016)
6. Song, J.S., Lee, S.W.: Automatic construction of positive/negative feature predicate dictionary for polarity classification of product reviews. J. KIISE: Softw. Appl. **38**(3), 157–168 (2013)

Design of Establishment System of Satisfaction Index for Tourist Sites According to the Weather Using Deep Neural Network

Hyeon-woo An[1], Kyungrog Kim[2], and Nammee Moon[1(✉)]

[1] Department of Computer Engineering, Hoseo University, Asan, Republic of Korea
grabroe@naver.com, nammee.moon@gmail.com
[2] Department of Electronic Engineering, Hoseo University, Asan, Republic of Korea
it4all@hoseo.edu

Abstract. Text's sentiment analysis techniques of texts play an important role in many fields including decision system. Sentiment analysis has been developed by combining with various technologies in accordance with this trend. The approach that extracts and uses the qualities in the sentences and the structure using the deep neural network belong to the example. In this paper, we propose a system to classify the vast amount of SNS tourist sites reviews into many factors that can affect tourism such as weather and includes rating by using the text sentiment analysis technology applying deep neural network.

Keywords: Sentiment analysis · Deep learning · CNN · LSTM

1 Introduction

Attempts have been made to ascertain the user's evaluation or perception of the product. From Paper questionnaire to Internet ratings to the latter, computer engineering techniques utilizing them have also made many advances. Examples might be a system that analyzes user ratings and the vocabulary frequency of reviews to derive product scores [1], or a method that analyzes review sentences using deep neural networks, such as the method used in this paper. However, the satisfaction with some tourist attractions depends on complex factors unlike products. Examples of factors include the convenience of tourism routes, the weather of the day, the concentration of tourists, and the degree of preservation of cultural assets in cultural tourism. One of the biggest factors may be weather [2]. Figure 1 is a picture of how climate, forecast, and Weather affect tourist's touring plans and tourism. Therefore, it is not a good way to recommend tourist attractions by averaging tourist evaluations in recommending them. That is, for example, tourist destinations that can look more beautiful on rainy days, and beaches with beautiful scenery such as waves and water are supposed to be high weighted to make recommendations on those days. However, for the above recommendations to be possible, a 'evaluation index' that is graded and classified by weather is required, and the data needed to build it must be very large. But according to the

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 310–315, 2020.
https://doi.org/10.1007/978-981-13-9341-9_53

research, less than 10 are non-popular tourist destinations in the case of the Tripadvisor and Google Maps review, which are platforms for managing reviews with ratings. This paper proposes a recommendation system that utilizes deep neural network and Instagram review data to compensate for the lack of data and to combine weather data using public data.

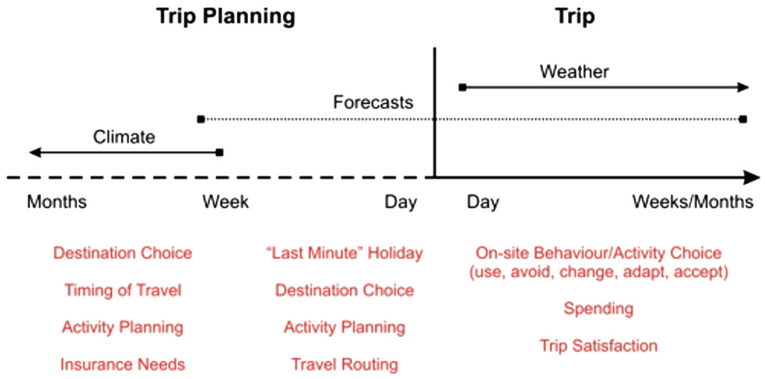


Fig. 1. Weather-climate information for tourist decision-making [2]

2 Body

The results of sentiment analysis of sentences are generally divided into positive/negative or five score from very positive to very negative. In this paper, it is necessary to apply the viewpoint of tourist divided by season and weather, so we will carry out the evaluation divided into 5 score. Since there is no dataset divided into five classes such as SST-1 in Korea, the rated review data will crawl rated reviews to be used as a dataset, with the target data being a Google Maps review and the Tripadvisor review based on user experience. This review data includes long text and ratings that are classified between 1 and 5 as score. The deep network structure that is responsible for the sentiment analysis of long texts uses a technology that combines CNN (Convolutional Neural Network) and LSTM (Long-Short Term Memory) [2].

In this paper, we propose a method for evaluating Korean sentences by preprocessing process that classify review data into morphemes and expressing them as numerical vectors, and a learning process for mapping preprocessed vectors to scores using deep learning. The preprocessing process is a process of classifying review sentences into morphemes and learning scores based on a vectorized array using word2idx. And then CNN and LSTM extract/preserve contextual information of words and learn in relation to ratings. Applying reviews as input to these learned systems would result in close ratings that fit the review within the error.

The data to be learned is an unproven, in other words a set of data that cannot be verified for similarity between score and sentences. Since there is currently no data available as a measure of reliability divided by ratings in the country, data for verification also uses collected review data.

There have been many studies on the use of deep learning in sentence analysis. Examples include positive/negative assessment using sentiment tree and RNN (Recursive) [3] and similarity determination between sentences using RNN and positive/negative evaluation combining CNN and RNN [4]. RNN is effective at learning variable and continuous data, and long-term dependencies that may exist in the long sentence can also be adequately addressed if an application uses technologies such as LSTM and GRU. CNN can also be used to extract and preserve the features of the context using its characteristic that can extract and preserve the local features [5].

3 Proposal System

The model presented in this paper can be divided into data collection, morpheme classification, sentence vectorization, and learning. In Korea, there is no verified Korean dataset with a rating, so it should be collected.

The proposed system is shown in Fig. 2. It uses weather data and tourist information data from Korea Meteorological Agency (KMA) and public data portal, which is used to combine weather information and tourist information on the day based on the name and date of the tourist sites included in the review data. As reviews with ratings, we can collect Tripadvisor and Google Maps Review to learn rating algorithm and add ratings to Instagram Review that does not include a rating. The resulting Rating Table contains ratings, tourist names, posting dates, and review sentences. When combined with the name of the tourist sites in the Rating Table and the public data called 'TURISM DATA', the coordinates of the tourist sites can be obtained. By using coordinates of the meteorological observation station provided by 'METEOROLOGICAL DATA', the stations closest to the tourist sites can be found, and by attaching using weather information in that station, can be spatial join for weather and tourism information. The final combined data are the evaluation index of the Fig. 2, a review table of tourist attractions, including weather and rating. As a result, if you use this table to get an average rating on a rainy day, you'll get a rating for each tourist attraction on a rainy day.

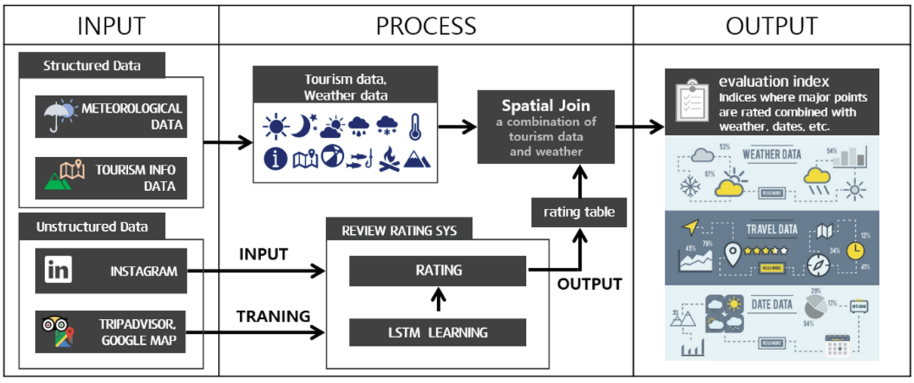


Fig. 2. Data used, data processing and output results (picture data source: Vecteezy.com)

3.1 Dataset

We chose Google Maps and TripAdvisor for the review platform with ratings to be used for learning. There are many other platforms, but mainly the reviews linked to the blog posts. Most of them are not suitable for learning data because they contain a large amount of articles which are not uniform. Also, since the criterion of rating is changed according to the platform characteristics, it is judged that the fusion of various platforms may lower accuracy. The list of tourist sites in Korea, which is collect to review, utilizes a list of about 5,000 tourist sites in the marine tourism information database provided by the public data portal.

We chose the review data of the Instagram as a dataset that will be rated by the learned system. In the case of Instagram, it has a characteristic that it records the user's daily life quickly. Therefore, it is judged that it is suitable for use as review data due to frequent posting on the day of travel and more data than other SNS platforms. As an example, whereas in Tripadvisor there are less than 40 reviews for 'the Gongseori church', in Instagram there are about 14,000 reviews.

Data collection is carried out using the Python Selenium library. The Fig. 3 shows the sample data collected from the Tripadvisor and Instagram for some tourist sites.

title	context	rating	date
굉장히 아	제목과 같이 굉장히 잘 꾸며서 아름답고 데이트	3	2017-12-08
아름다운	아산 스파비스에서 멀지 않습니다. 잘 가꿔진 정	5	2017-11-11
개인 소유	식물원이에요. 개인 소유의 장소라서 더욱 관리가 잘 되고 있는 느낌. 늘 깔끔하게 정돈되고 청결한 곳이라서 유아동반 가족들 봄나들이로 아주 좋아요. 지금 시즌에 가기 정말 좋은 곳입니다.입장료도 있고요.	5	2017-03-24
꽃을 느낄	외도 보타니아 설립자의 자제 분이 하는 걸로 알	5	2017-03-23
연인들이	어린 아이들 데리고 갔었어요. 가벼운 등산 겸	4	2016-10-26
잘 정돈된	잘 정돈된 정원에서 데이트를 즐길 수 있는 곳(	4	2016-01-27

2014년 08월 19일	피나클랜드	#휴가#피나클랜드#알맹#베프 마음이 트이는곳, 날씨가 좋았다면 더 좋았을걸 ! 추억 1+1😢
2014년 08월 16일	피나클랜드	#빵실빵실#휴가기간#얼굴 터질라 #아산#피나클랜드 적당한 기온 적당한 시간 딱 좋았다
2014년 08월 15일	피나클랜드	#아산#피나클랜드 생각보다 좋은데? 굿굿
2014년 07월 27일	피나클랜드	아~조티♥ #피나클랜드 #여름#수목원#꽃 #너무좋아
2014년 07월 26일	피나클랜드	이름모르는 #꽃 #예뻐 #여행 하기조은 날씨✪ 여기는 #피나클랜드

Fig. 3. (left) Crawled Tripadvisor Sample Data, (right) Crawled Instagram Sample Data

3.2 Preprocess

Classification of morphemes is the process of classifying Korean sentences into smaller units suitable for learning. For example, if there is a review sentence, "잘 보존된 좋은 장소입니다." classification such as (잘/MAG 보존/NNG 되/XSV ㄴ/ETM 좋/VA 은/ETM 장소/NNG 이/VCP ㅂ니다/EF) enables vectorization to be efficient for learning. When the classification of the morpheme for the review sentence is completed, word embedding is carried out through word2vec.

3.3 CNN and LSTM

The learning model used in this paper cites already published research results [4]. CNN is a method that is often used to classify images by extracting and preserving local features of images composed of matrix data. The feature extracting characteristics of

this method can be applied to extract the contextual features existing in the sentence [5]. In this case, LSTM performs sentence encoding using results extracted via CNN and draws ratings through fully connected softmax layer.

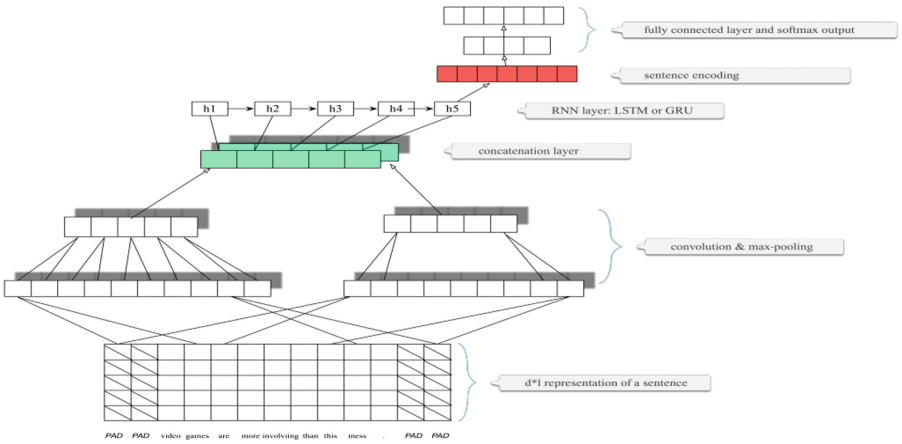


Fig. 4. Model architecture for an example sentence [4]

The Fig. 4 proceeds as follows. The matrix at below is a k*n matrix, each of which is embedded into a k-dimensional vector by the pre-processing process, which is convolution by window of a given size and feature maps of the number of convolutional kernels are generated. At here, the task of extracting and preserving the local feature of the sentence occurs. Here we proceed with the window size of 4 and 5 and m convolution. When max pooling is performed on feature maps of different sizes, $m \times 2$ feature maps are generated, and they are concatenated for transfer them to LSTM.

$$Z = \oplus(P_4, P_5), Z \in \mathbb{R}^{\left[\frac{l-w+1}{2}\right] \times (m \times 2)} \tag{1}$$

The output from CNN is like word-based N-grams. The LSTM then produces the same output as the entire sentence encoded by the input word matrix. Then, the LSTM can form a penultimate layer as a result and obtain a probability distribution through a fully connected softmax layer. The softmax operation is calculated as follows:

$$\hat{P}_i = \frac{\exp(o_i)}{\sum_{j=1}^{C} \exp(o_j)} \tag{2}$$

The formula below shows cross entropy as a loss function, where T is the training corpora, V is the number of categories (5 score), and P is the V-dimensional one-hot coded actual training data with 1 and 0. The entire model is trained end-to-end with stochastic gradient descent [4].

$$\text{loss} = -\sum_{s \in T} \sum_{i=1}^{V} \hat{P}_i^t(C) \log\left(\hat{P}_i(C)\right) \tag{3}$$

4 Conclusion

If it is possible to evaluate tourism sites by applying various environmental factors such as time, season, and weather, it can be used as important information for tourism industry such as improvement of tourist sites and recommendation system. However, it was determined that the platforms that have accumulated tourist review did not have enough data to analyze the above. So, we wanted to propose a system that even did not include the rating but would utilize a much wider platform for reviewing tourist attractions.

However, due to the characteristics of the SNS data, there is a possibility that the time when the user posts the article and the time when the visitor visits the tourist sites are not the same. Since it is also prohibited to collect metadata such as exif of photographs in the Instagram, it is judged that the process using social engineering will be necessary for perfect application.

Acknowledgments. This work has supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2016R1C1B1015499, NRF-2017R1A2B4008886).

References

1. Scaffidi, C., et al.: Red Opal: product-feature scoring from reviews. In: Proceedings of the 8th ACM Conference on Electronic Commerce, pp. 182–191. ACM (2007)
2. Scott, D., Lemieux, C.: Weather and climate information for tourism. Proced. Environ. Sci. **1**, 146–183 (2010)
3. Socher, R., et al.: Recursive deep models for semantic compositionality over a sentiment treebank. In: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, pp. 1631–1642 (2013)
4. Wang, X., Jiang, W., Luo, Z.: Combination of convolutional and recurrent neural network for sentiment analysis of short texts. In: Proceedings of COLING (2016), the 26th International Conference on Computational Linguistics: Technical Papers, pp. 2428–2437 (2016)
5. Kim, Y.: Convolutional neural networks for sentence classification. arXiv preprint arXiv: 1408.5882 (2014)

A Study on PGP (Pretty Good Privacy) Using Blockchain

Chang-Hyun Roh and Im-Yeong Lee(✉)

Department of Computer Software Engineering, Soonchunhyang University,
22, Soonchunhyang-ro, Sinchang-myeon, Asan-si
Chungcheongnam-do, Republic of Korea
{rohch, Imylee}@sch.ac.kr

Abstract. E-mail has developed as the Internet has evolved. Encryptions such as Secure/Multipurpose Internet Mail Extensions or Pretty Good Privacy (PGP) are used to ensure privacy and confidentiality. However, the public PGP key is stored on a server and is synchronized with the PGPs of other servers, and thus is not in real time. In addition, public key search, registration, and cancellation are difficult when a key server is down. To solve this problem, we develop a key management server integrating PGP into a blockchain; this allows all participating servers to synchronize registration and deletion in real time.

Keywords: Blockchain · Distribute ledger technology · Pretty Good Privacy

1 Introduction

E-mail has evolved along with the Internet. The use of a Web browser to access e-mail has become routine, and technical perfection has been attained. Recent developments in Social Network Services (SNSs) and instant messaging have greatly reduced personal usage, but many businesses continue to rely greatly on e-mail [1]. A great deal of effort has been devoted to e-mail confidentiality. E-mail security requires protected communication and encryption of the e-mail content [2]. Network security systems include the Secure HTTP and the Secure Socket Layer (SSL) methods. Secure/Multipurpose Internet Mail Extensions (S/MIMEs) and Pretty Good Privacy (PGP) are commonly used to encrypt e-mails. PGP affords high-level security [3] using public key infrastructure (PKI)-based algorithms to ensure e-mail confidentiality, integrity, and authentication by reference to the public key of the recipient. However, a key server manages public keys. A public key management system synchronizes the operations of servers in many organizations; synchronization is not performed in real time. If a key management server is down, key discovery, registration, and cancellation become very difficult. Thus, single-point-failure is in play. The synchronization and single-point error problems associated with PGP can be solved using blockchain techniques. Blockchain was first featured in the peer-to-peer (P2P) cryptography of Distributed Ledger Technology; integrity is assured by connecting all blocks containing the same data to their prior and subsequent blocks [4].

Here, we describe a key management server that can be implemented within a blockchain network to solve the abovementioned PGP problems.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 316–320, 2020.
https://doi.org/10.1007/978-981-13-9341-9_54

2 Related Research

2.1 PGP

PGP was created by Phil Zimmermann in 1991. PGP encrypts e-mails and associated files, ensuring both confidentiality and authentication. It has been repeatedly refined, and in 2001 was designated an encryption standard by the Internet Engineering Task Force (IETF) [5].

PGP features an encryption algorithm (Fig. 1) employing both symmetric and public key encryption. When encrypting a message, a session key is randomly generated using a symmetric key cryptosystem [6]. The public key cryptosystem encrypts the session key together with the recipient's public key, then encrypts the message, and sends the session key to the recipient. The receiver decrypts the session key using a private key, and then decrypts the message employing the newly accessible session key. Table 1 shows the relevant encryption algorithms. The session key is generated by the user, but the public key must be obtained from the recipient, either directly or via registration with a public key server. Most users access online public key servers to retrieve recipient identities and obtain the public keys required for PGP transfer.

2.2 Blockchain

Blockchain is a cryptocurrency created by Nakamoto Satoshi in 2009. Bitcoin users transfer currency and engage in transactions that are not mediated by institutions such as central banks. In a P2P network, all participants record identical transactional data; they cooperate to maintain the network. In a blockchain, all participants access identical data in block form; the blocks are associated with hash values that prevent tampering. Blockchains may be public, private, or consortium-based. Anyone can create a block by joining a public blockchain. Any person or consortium may seek to create a network, but only authorized participants or organizations can create/validate the block. The blockchain structure is shown in Fig. 1.

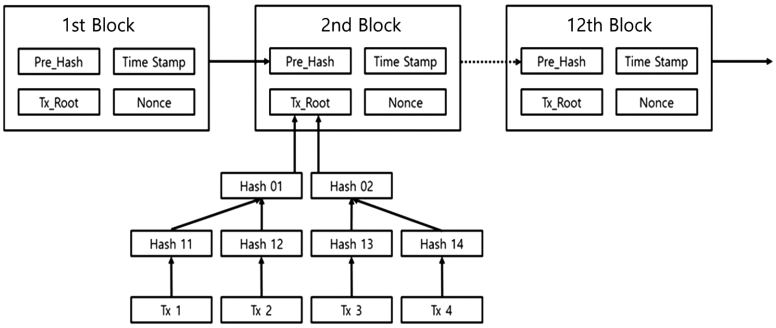


Fig. 1. Blockchain structure

3 Proposed Scheme

The proposed method can perform key registration, key stop, key signing, and key retrieval functions in a blockchain network. The overall scenario of the proposed Scheme is shown in Fig. 2.

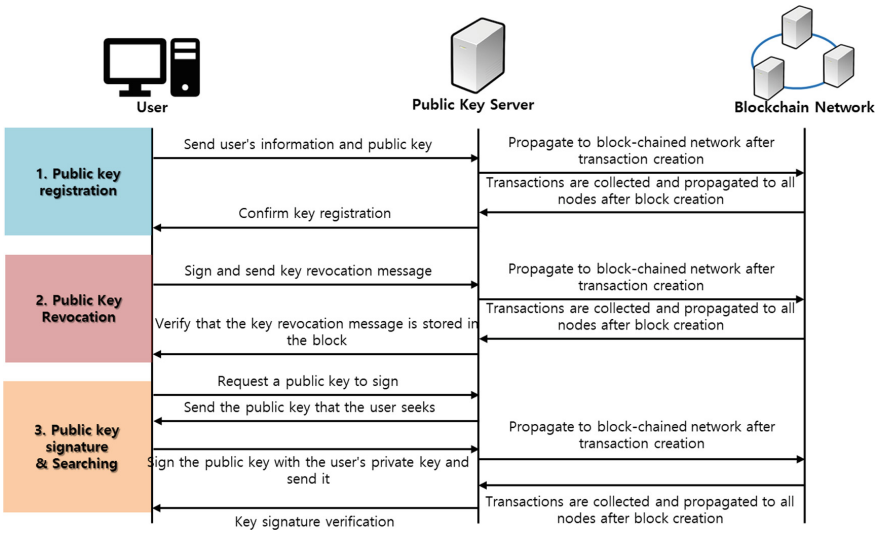


Fig. 2. Overall scenarios of the proposed scheme

3.1 System Parameters

Table 1. System parameters

Parameters	Define	Parameters	Define
PK_U	User U's public key	SK_U	User U's private key
$Info_U$	User U's information	$Block_1$	i - th block
Sig_{SK_U}	Signed with private key	TS	Transaction
$R \cdot Cert$	Public key revocation certificate		

*: participate object (U: User, PKS: Public Key Server)

3.2 Public Key Registration

The user generates a public/private key pair, sends the user information and public key to the public key server, and the public key server creates and propagates the transaction. The blockchain network completes the registration of the public key by

generating a block through the consensus algorithm and connecting it to the blockchain through validation and block propagation.

1. U generates $PK_U \cdot SK_U$
2. U sends PK_U and $Info_U$ to PKS
3. PKS generates PK_U and $Info_U$ as TS
4. PKS deploys TS as a blockchain network
5. Blockchain network collects TS every minute
6. Collect the collected TS and create a $Block_1$
7. $Block_1$ identification and propagation to blockchain networks
8. PKS sends the header of the generated block to U

3.3 Public Key Revocation

Public keying sends a revocation certificate to the public key server to prevent the user from using the stored public key. public key server creates a transaction with the revoked certificate and forwards it to the blockchain network. blockchain network then creates a block and links it to the block to disable the key.

1. $R \cdot Cert_U$ is created and signed with private key: $Sig_{SK_U}(R \cdot Cert_U)$
2. $R \cdot Cert_U$, $Sig_{SK_U}(R \cdot Cert_U)$ are send to PKS
3. PKS validates $Sig_{SK_U}(R \cdot Cert_U)$ as PK_U
4. Generate TS by inserting $R \cdot Cert_U$ after validation
5. Blockchain network collects TS every minute
6. Collect the collected TS and create a $Block_1$
7. $Block_1$ identification and propagation to blockchain networks
8. PKS sends the header of the generated block to the U
9. Confirm that the public key has been revoked

3.4 Public Key Signature

Public key signatures occur when you trust another user's public key in the public key ring. Users sign their public key for other trusted users to indicate their trust.

1. U determines the PK_1 to sign and generates $Sig_{SK_U}(PK_i)$
2. PK_1, $Sig_{SK_U}(PK_i)$ are send to PKS
3. PKS validates $Sig_{SK_U}(PK_i)$ as PK_1
4. PKS checks for duplicate $Sig_{SK_U}(PK_i)$ in the *Block*.
5. If duplication does not occur, insert PK_1 and $Sig_{SK_U}(PK_i)$ into TS, generate and propagate to blockchain network
6. Blockchain network collects TS every minute
7. Collect the collected TS and create a $Block_1$
8. $Block_1$ identification and propagation to blockchain networks

3.5 Public Key Search

Public key search is used when a user adds another user to the public key ring. The user uses the data of another user (user ID, key ID, e-mail address, etc.) to request a search to the public key server.

1. U learns $Info_1$ of another user
2. U to search the PKS for another user's information $Info_1$
3. PKS retrieves $Info_1$ from $Block$
4. Reply PK_1 and $Sig_{SK_U}(PK_i)$ to U if successful

4 Conclusion and Future Research

In this paper, we propose a public key server composed of blockchain network in PGP public key server. The public key server operated by the agency had a single point of failure that could not provide service when the server was stopped due to an attack. In addition, public key servers operated by each organization do not provide real-time synchronization, so you cannot store keys securely.

The disadvantage of existing public key servers is solved by configuring a blockchain network to solve a single point of failure. We also proposed a distributed system that provides data integrity because all servers are synchronized in real time. In the future, the proposed technique should be studied in such a way that users' public keychain management can be automatically added to smart contracts.

References

1. Radicati Group: Eamil Statistics Report, 2017–2021. Radicati Group, 6 February 2017
2. Park, D.-U., Park, J.-H., Kim, J.-S., Kim, I.-M.: A web based secure e-mail system using the PGP algorithm. KIPS Trans. Part C **8C**(1), 16–22 (2001)
3. Rhee, M.-Y., Kim, J.-H., Ryou, J.-C., Song, Y.-J., Youm, H.-Y., Lee, I.-Y.: E-Commerce Security Technology. Saengneung Publisher (1999)
4. Nakamoto. S.: Bitcoin: A Peer-to-Peer Electronic Cash System (2008). http://bitcoin.org
5. IETF: MIME Security with OpenPGP. RFC3156, August 2001
6. Stallings W.: Cryptography and Network Security: Principles and Practice, Global Edition, Pearson Education Limited (2016)

Mining Semantic Tags in a Content Analysis System for a Letter Database of Ethnic Koreans Living in China

Hyon Hee Kim[1], Yuntae Kim[1], Hyong-Jin Moon[1], Jungsun Choi[1], Joo-mi Lee[1], Soungho Ye[2], Seunghyun Seo[1], and Jinnam Jo[1(✉)]

[1] Dongduk Women's University, 60 Hwarang-ro, 13-gil, Seongbuk-Gu, Seoul, South Korea
{heekim, runtai, subugwil2, jschoi, chanbi, seoseung4, jinnam}@dongduk.ac.kr
[2] Seoul University of Foreign Studies, 85 Nambusunhwan-ro 356-gil, Seocho-Gu, Seoul, South Korea
ruishenghao@gmail.com

Abstract. In this paper, we present a content analysis system for the letter database for ethnic Koreans living in China using several machine learning techniques. First, letters from ethnic Koreans living in China to Korea Broadcasting System (KBS) to find separated families were digitized and constructed as a letter database. Second, to help retrieval of digitized letters, semantic tags were annotated to each letter manually. Third, to find the contents of the digitized letters, machine learning techniques were applied to those semantic tags. Tag-based topic modeling was performed and a tag network was constructed. Over 150,000 letters were scanned as image files and have been constructed as a letter database. Among them, 3,500 letters are randomly selected and analyzed. Our approach shows that the semantic tags play an important role in the digitized letter retrieval. We are currently enhancing algorithms for tag analysis, and the developed system will broaden the horizon of digital humanities.

Keywords: Content analysis · Tag-based topic modeling · Mining semantic tags · Letter retrieval using tags · Letter database of ethnic Koreans in China

1 Introduction

Digital humanities (DH) has attracted much attention as the crossroads of information technologies and humanities. Although the precise definition of digital humanities is still floating, it is generally defined as "Along with the digital archives, quantitative analyses, and tool-building projects that once characterized the field of humanities" [1]. One of the well-known DH project is "Mapping the republic of letters" undergoing at Stanford University [2]. In this project, the letters of famous intellectuals in Europe and America in the 17th and 18th centuries were collected, digitized, and analyzed. Their research showed not only correspondence metadata and correspondence network but also perspectives of contemporary intellectuals.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 321–326, 2020.
https://doi.org/10.1007/978-981-13-9341-9_55

In this paper, we present the ongoing DH project at Dongduk Women's University. Over 150,000 letters sent from the ethnic Koreans in China to a radio program of Korea Broadcasting System (KBS) to find their separated families from 1974 to 1992 are digitized and have been constructed as a database system. To analyze contents of correspondences, semantic tags describing the contents are annotated to each letter manually. The tags are used for two purposes. One is summarizing contents of the letters, and the other is helping a letter retrieval by end users. The content analysis system was developed for mining the annotated tags. It provides the visualization of important semantic tags which will be used for the letter retrieval.

The contribution of this paper is in the development and validation of the semantic tag-based framework for constructing letter databases and analyzing their contents. For this purpose, 3,500 letters are randomly selected from the letter database and analyzed their contents using the topic modeling and the tag network. The results of topic modeling are used as the definition of categories and subcategories for the letters, while the results of the tag network are used to extract important semantic tags.

The remainder of the paper is organized as follows. In Sect. 2, we present an overview of the system. In Sect. 3, we explain the mining semantic tags in detail. Finally, in Sect. 4, we give concluding remarks.

2 System Overview

Figure 1 shows the overview of the semantic tag-based framework for the database construction and the content analysis. The digitized letters are inserted into the database with semantic tags, and then uploaded on the database system. In this step, Korean-Chinese dictionary is generated for the content analysis and the letter retrieval shown in bottom left side of Fig. 1. Then, the content analysis is performed using the content analysis engine. Frequently annotated semantic tags are visualized and a word cloud is used as search keywords.

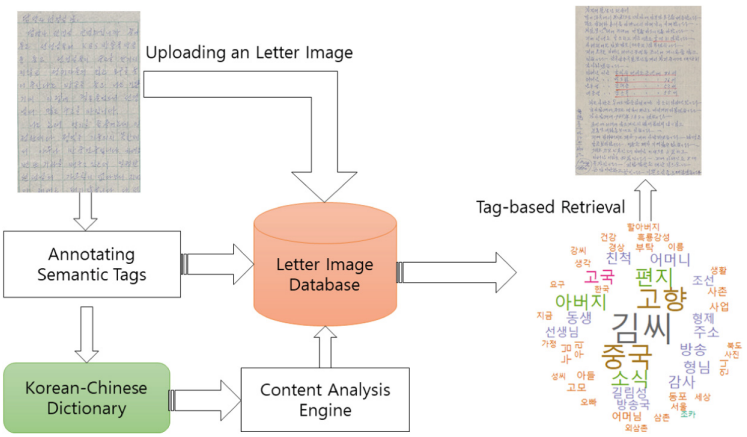


Fig. 1. System overview

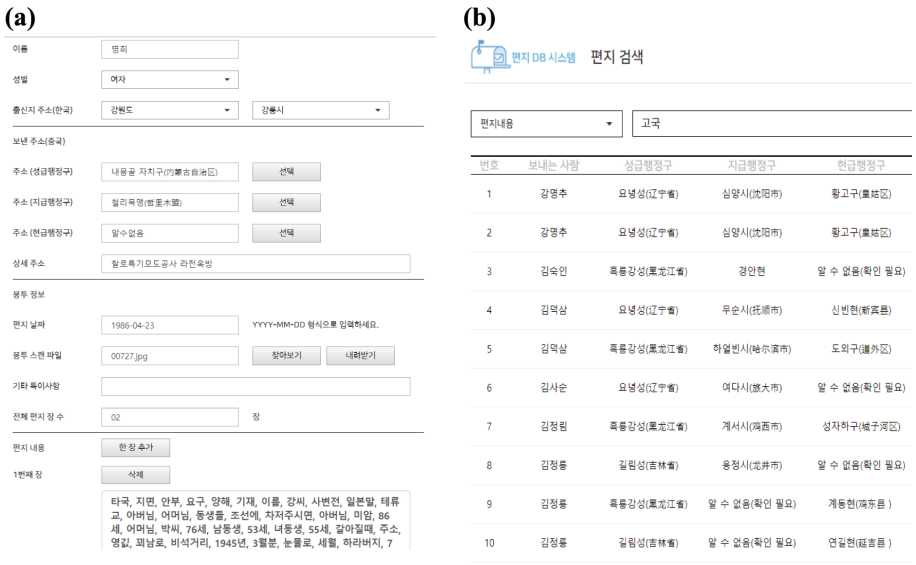


Fig. 2. (a) User interface of the system (b) Search results

The Fig. 2(a) shows the user interface of the developed system. The sender information, like names, genders and past home addresses in Korea and current addresses in China, dates of the letter, semantic tags describing contents are inserted into the database system. The system provides two-layer retrieval. One is using the semantic tags and the other is using the predefined categories and keywords. The categories and keywords are defined for further research obtained from diverse research area such as literature appearing in the letter, languages of ethnic Koreans in China, politics and economy in China, etc. The categories are still being updated considering the results of the topic modeling. The Fig. 2(b) shows that search results with keyword "homeland". The system was developed by MS SQL Server for the database and R for the content analysis in the windows environment.

3 Mining Semantic Tags

Tag-based image retrieval is widely applied to research area of information retrieval [3, 4]. Differently from other tag-based image retrieval system, since characters are written in both Korean and Chinese with cursive letters sometimes in our system, it is much more difficult to recognize contents from the letters using automatic image recognition technique than other general images. Therefore, semantic tags play an important role in retrieval of letters. In particular, for mining the semantic tags, Korean-Chinese dictionary is semi-automatically generated and used for content analysis process. For the content analysis, the tag network analysis developed in our former research [5] and topic modeling [6] are used.

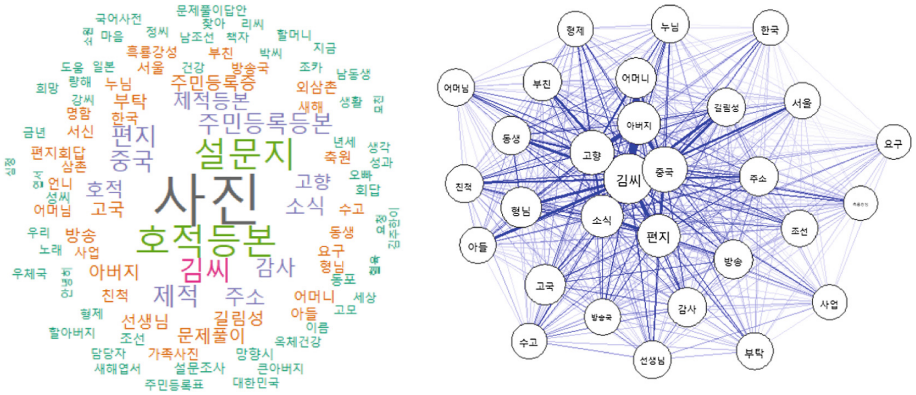


Fig. 3. Visualization of content analysis

Figure 3 shows the visualization of content analysis. The left side of the Fig. 3 is visualization of top 100 important tags based on Term Frequency-Inverse Document Frequency (TF-IDF) [7]. The most important tags contain photo, family register, identification card which we have attributed to the fact that the letters were used to find separated families. Hometown, address, named entity for family such as mother, father, cousin, uncle, etc. are highly ranked.

The right side of the Fig. 3 is visualization of the tag network with the top 30 frequent tags. In the tag network, each tag is represented as node, and relationships between tags are represented as edges. If annotated tags are shared in different letter, the tags are connected via edges. The most closely connected tags are Kim (one of the most common family name in Korea), hometown, news and thanks.

To analyze the contents of the letter more specifically, the 8 categories are defined by professionals in each field. The 8 categories are as follows: migration, life, folk, literature, town, family, politics, and economy. Each category has subcategories, and the categories are used to classify the letters according to the subjects. The migration category contains subcategories such as migration route, transportation for migration, and location and time of settlement, etc. The life category contains income, occupation, education, hobby, health, and marriage. The folk category contains food, clothing, and shelter, while the literature category contains poems, novels, their own poems, etc. The categories are changeable by integrating the topic modeling results and current categories. A letter may be classified maximal three categories.

The category-based analysis is also performed. For example, in the category of politics, national unification, development, peaceful unification, north Korea, and motherland are selected as important tags, which means desire and wishes of ethnic Koreans in China. In the category of economy, Korean dictionary, Korean-Japanese dictionary, book, magazines, tapes are highly ranked, and it can be inferred that senders of the letters request their needed goods. Other category-based analysis is also in progress.

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
[1,]	"편지"	"형님"	"김씨"	"김씨"	"고국"
[2,]	"고국"	"김씨"	"중국"	"고향"	"고향"
[3,]	"감사"	"성씨"	"고모"	"중국"	"동포"
[4,]	"중국"	"누님"	"아버지"	"북도"	"감사"
[5,]	"책자"	"편지"	"할아버지"	"경상"	"발전"
[6,]	"부탁"	"중국"	"고향"	"소식"	"선생님"
[7,]	"방송"	"소식"	"삼촌"	"동생"	"생각"
[8,]	"학습"	"사촌"	"소식"	"언니"	"마음"
[9,]	"요구"	"동생"	"길림성"	"길림성"	"사진"
[10,]	"일본어"	"주소"	"할머니"	"어머님"	"민족"

	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
[1,]	"중국"	"어머니"	"방송"	"문제"	"김씨"
[2,]	"편지"	"아버지"	"고국"	"김00"	"중국"
[3,]	"소식"	"김씨"	"선생님"	"고향"	"고국"
[4,]	"아버지"	"외삼촌"	"건강"	"吏<84>"	"편지"
[5,]	"고향"	"중국"	"편지"	"풀이"	"조선"
[6,]	"우리"	"고향"	"요구"	"감사"	"어머님"
[7,]	"지금"	"오빠"	"감사"	"일어"	"흑룡강성"
[8,]	"감사"	"언니"	"옥체"	"편지"	"소식"
[9,]	"주소"	"강씨"	"축원"	"노래"	"주소"
[10,]	"서울"	"동생"	"동포"	"한국"	"부탁"

Fig. 4. Result of topic modeling

Figure 4 shows the result of topic modeling with 10 topics and 10 tags. The purpose of topic modeling is providing a guideline to classify the letters into diverse categories. 10 tags in topic 1 are letter, hometown, thanks, book, learning, Japanese, request, favor, China, broadcasting, which represents requests for goods. Topic 5 contains homeland, hometown, compatriots, thanks, development, people, mind, think, teacher, photo, and describes homeland. Also, topic 7 represents family with 10 tags such as mother, father, uncle, brother, sister, etc. To look for the appropriate topics, diverse experiments with different topics and tags are performed, which will be integrated into the definition of categories.

4 Conclusions

In this paper, we present a semantic tag-based letter retrieval and analysis system. The 150,000 letters from China to KBS radio program to find separated families were digitized and have been constructed a letter image database. And then the content analysis is performed on the database. First, important tags are selected using TF-IDF weight and used for letter image retrieval. Second, the tag network is constructed to find closely related tags. Finally, topic modeling is performed to classify the letters based on the predefined categories.

The research presented in this paper is the ongoing project and randomly selected 3,500 letters among 150,000 letters are analyzed. Current results of content analysis show that contents of letters are very limited to finding separated families and request of the needed goods. However, we expect that if the whole letters are analyzed, more diverse topics describing living cultures of ethnic Koreans in China can be found. We are currently developing an algorithm looking for important tags using the tag-network. As a future work, we plan to develop a visualization tool for geospatial data such as hometown addresses and residential addresses.

Acknowledgments. This work was supported by Korean Studies Foundation Research through the Ministry of Education of the Republic of Korea and Korean Studies Promotion service of the Academy of Korean Studies (ASK-2017-KFR-1230011).

References

1. Gardiner, E., Musto, R.G.: The Digital Humanities. Cambridge University Press, Cambridge (2015)
2. Mapping the Republic of Letters. http://republicofletters.stanford.edu/Lavretsky
3. Chen, L., et al.: Tag-based image retrieval improved by augmented features and group-based refinement. IEEE Trans. Multimed. **14**(4), 1057–1067 (2012)
4. Lu, D., et al.: Tag-based image search by social re-ranking. IEEE Trans. Multimed. **18**(8), 1628–1639 (2016)
5. Kim, H.H., Kim, D., Jo, J.: Patent data analysis using clique analysis in a keyword network. J. Korean Data Inform. Sci. Soc. **27**(5), 1–12 (2016)
6. Kim, H.H., Rhee, H.Y.: Trend analysis of data mining research using topic network analysis. J. Korea Soc. Comput. Inform. **21**(5), 141–148 (2016)
7. Salton, G., McGil, M.J.: Introduction to Modern Information Retrieval. MacGraw-Hill, New York (1986)

A Suggestion of Image's Position Analysis System Using Beacon Positioning

Seo Kyung Jung, Sung Geun Yoo, and Sangil Park(✉)

Nano IT Design Fusion Graduate School, Seoultech, 232 Gongneungno, Nowon-Gu, Seoul 01811, Republic of Korea
tjrudl438@naver.com, orcogre@gmail.com, sipark@seoultech.ac.kr

Abstract. When a 360° image is produced from images taken by a plurality of general cameras, an image stitching technique for finding the same area between a plurality of images and correcting the geometry is required. The image stitching technique compares the similarities of feature points extracted from several images, and connects the images together to form a large image. In this paper, assuming a theater with a beacon attached, the position of the photographer is roughly grasped through the beacon, and the position of the photographer captured by the user through the application, the azimuth angle of the image and the field of view (FOV). We propose a method of filtering images with stitching object images.

Keywords: Metadata · Multi camera video stitiching

1 Introduction

Currently, most high-quality 360-shots are made with multiple cameras fixed on one mount. The conventional method is that it is difficult to produce contents because a distortion of a wide angle lens can not be avoided because an image is taken with a small number of cameras equipped with a wide angle lens and an expensive device is used. [1] In order to overcome the limitation of the production environment, methods of producing 360 images by shooting with a large number of general cameras are being developed. In order to perform image stitching, it is necessary to find the same area between a plurality of images and to perform geometric correction. Extract feature points from the image using feature extraction algorithm and find corresponding points between images. However, there is a difficulty in finding images that can stitch images directly captured by a plurality of users other than the cameras fixed in the league in the room. Conventional methods are grouping based on camera sensor data to reduce the time required to find the correspondence points between outdoor and outdoor images [2], and increasing the speed of stitching processing using inertial sensor information [3] If the image is taken outdoors, the GPS (Global Positioning System) information can be extracted from the EXIF (Exchangeable Image File Format) information of the image to know the photographing position information. However, Line Of Sight is not ensured, and thus there is a disadvantage in that the received signal is very weak in the indoor positioning. [4] Various indoor positioning technologies have been developed to

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 327–331, 2020.
https://doi.org/10.1007/978-981-13-9341-9_56

overcome these limitations. In this paper, we use a representative BLE (Bluetooth Low Energy) based indoor positioning technology to obtain the location of indoor images, and using the azimuth and FOV (field of view) information obtained from the camera sensor. We propose a method to filter by neighboring images as much as possible. The video with indoor location information and extracted metadata are stored in the database through the filtering algorithm based on the metadata that is transmitted to the server and obtained. Thereby sorting the images that are likely to stitch. In this way, it is expected that the error and complexity of the stitching using multiple cameras can be reduced.

This paper is composed as follows. In Sect. 2, beacon-based positioning techniques and azimuths are described. Section 3 explains the outline and detailed description of the proposed application and system and concludes in Sect. 4.

2 Related Works

(1) Beacon
The beacon is a communication method using the Advertise mode used in the BLE (Bluetooth Low Energy) mode. The BLE device transmits an advertiseing packet to any nearby terminals, The packet includes data such as a device ID, a service place name, and a transmitted signal strength. The receiving terminal computes the position using the ID of the BLE device or the distance between the receiving terminal and the BLE device using the reference point position or the multifunctional position [5]. This beacon-based positioning technology has a transmission distance of up to 50 m compared with NFC, allowing positioning alone, but it can produce synergy by combining with existing positioning technology. In the case of a beacon, the distance between the beacon and the mobile is obtained from the RSSI (Recieved Signal Strength Indicator) for distance measurement. The signal intensity of the beacon and the received terminal is measured and used, and the distance is accumulated using the signal strength loss or the path loss between the terminal and the beacon.

(2) The azimuth angle
The azimuth angle indicates the horizontal angle in the clockwise direction relative to the true north. It is expressed from 0° to 360° and tells the direction of the camera at the time of shooting. Even if an image is photographed at the same place in a theater with a beacon, azimuth information is essential because the image can not be stitched if the shooting direction is different. The azimuth is measured using the Android geomagnetic sensor.

(3) The traditional Stitching method
Input. n ordered image sequence $S(S0, S1,\ldots, Sn-1)$ which mutually have overlapping areas:
   (1) The first two images are selected in the set S to extract the SIFT feature points.
   (2) In the S sequence, designate the first image as the reference and the second image as the new input image. KNN (K Nearest-Neighbors) algorithm is used to find the matching point of two images matching Euclidean distance.

(3) Depending on the matching point data, we use the RANSAC algorithm to find the affine matric H. The affine matric H transforms the new image into the space of the existing reference.
(4) Use the L-M algorithm to optimize the affine matrix.
(5) Transform a new image through H matrix.
(6) Find the optimal seam between the transformation result and the reference image, and stitch the images by combining along the optimal seam.
(7) Replace the first input image and the reference image in set S with the stitched image. In the next stitching process, the stitched image is selected as the reference image.
(8) Returning to the first step, the stitching process is performed, and the process continues until there is only one stitched panorama image in the set S [5].

3 Proposed Method

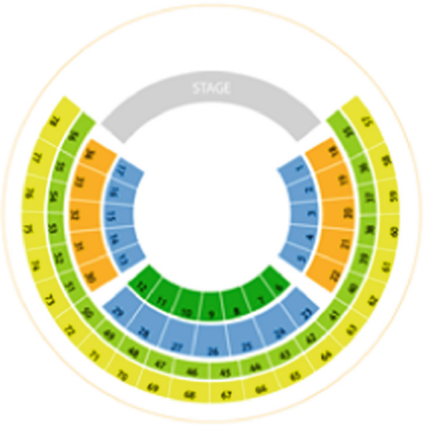


Fig. 1 Zones divided in the concert hall

In this chapter, we propose a method to filter images that can be stitched among images uploaded to the server based on the techniques described in Sect. 2. The proposed system consists of user's smartphone, server, and database. Smartphones are used for video, photo shoots and video is uploaded through mobile applications. The server has a location information database for the venue area, classifies the uploaded images by zones, and performs an algorithm for filtering based on azimuth and FOV information. Stitch processing time can be reduced by stitching the finally filtered images in the area.

3.1 How to Measure Beacon Signal

We divide the inside of the concert hall where we want to serve and divide the images of each area. First, several transmitting BLE devices are installed in the concert hall.

Since a large number of beacons are installed in one place, it is recommended to use the same UUID in advertising packets. Set the Major and Minor values to distinct areas within the venue. The terminal receiving the beacon signal in the corresponding area calculates the position using the reference point position or the multifocal position using the distance between the beacons. The calculated position information is increased in accuracy using a Kalman filter and then stored in the mobile (Fig. 2).

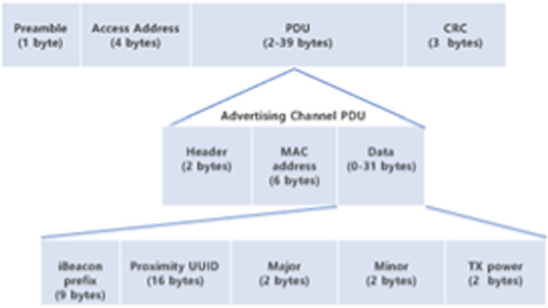


Fig. 2 Packet structure of iBeacon

3.2 Filtering Adjacent Images in the Database

After confirming that the uploaded image is within the zone range based on the position data, the uploaded image is stored in a specific zone (e.g., zone 10 in Fig. 1) in the database. After filtering the images taken from the adjacent places first. In the second lane, filtering is performed based on the azimuth information and FOV information of the photo in the meta data. Even in the same area, it is possible to reduce the number of images to be stitched by removing images taken in different directions. Stitching is performed using the SURF stitching algorithm for images that are filtered in detail in a specific area.

4 Conclusion

We proposed a method of filtering the image for 360 image stitching based on multi camera based on azimuth and FOV information obtained from the camera and measuring the position of the user through installed beacons in the venue. Conventional image stitching is performed by extracting feature points of images and comparing similarities between extracted feature points. It is expected that stitching can be processed more efficiently by reducing the number of feature point comparison objects through the proposed system in this paper. After that, it is expected to be used in various fields such as augmented reality, SNS connection service, as well as stitching using multiple cameras

Acknowledgement. This work was supported by 2018 Project to Support Leading Model for Commercialization with a University (Implementation of the Academic-Educational Cooperation System based on the Academic Linkage through the Establishment of Hongneung Tech-Biz OPERA Center)

References

1. Jeong, J., Jun, K.: High resolution 360 degree video generation system using multiple cameras. J. Korea Multimedia Soc. **19**(8), 1329–1336 (2016)
2. Im, J., Lee, E., Kim, H., Kim, K.: Images grouping technology based on camera sensors for efficient stitching of multiple images. J. Broadcast Eng. **22**(6), 713–723 (2017)
3. Kim, M., Kim, S.-K.: Enhancement on 3 DoF image stitching using inertia sensor data. J. Broadcast Eng. **22**(1), 51–61 (2017)
4. Seo, B., Jeong, E., Kim, N., Jang, J., You, D., Kim, D.H.: Image transform method of 360VR image acquired from six cameras for real-time streaming video. J. Broadcast Eng., 51–52 (2016)
5. Qu, Z., Lin, S.P., Ju, F.R., Liu, L.: The improved algorithm of fast panorama stitching for image sequence and reducing the distortion errors. Math. Prob. Eng. **2015**, 12 (2015). https://doi.org/10.1155/2015/428076. Article ID 428076

Anti-motion Blur Method Using Conditional Adversarial Networks

Myeong Gyu Lee[1], ChengNan Lu[1], Daniel Chung[3], Wee Jia Foon[4], Ilju Ko[2], Kok Yoong Lim[4], and Jinho Park[2(✉)]

[1] Department of Digital Media, Soongsil University, 369 Sangdo-Ro, Dongjak-Gu, Seoul 156-743, Korea
{brstar96, ryuck}@soongsil.ac.kr
[2] Global School of Media, Soongsil University, 369 Sangdo-Ro, Dongjak-Gu, Seoul 156-743, Korea
c2alpha@ssu.ac.kr
[3] Department of ICMC (Information Communication Materials, and Chemistry) Convergence Technology, Soongsil University, 369 Sangdo-Ro, Dongjak-Gu, Seoul 156-743, Korea
danielc@soongsil.ac.kr
[4] Creative Multimedia, Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Selangor, Malaysia

Abstract. In general, taking a high-speed object such as projectiles with a low-speed camera are accompanied by artifacts called so-called Motion blur. Motion blur is a phenomenon that the boundaries of a moving object diffuse unclearly. Motion blur is divided into a 'captured motion blur' and 'display motion blur'. The formal occurs when the object moves faster than the camera shutter speed, and the later occurs due to the limitations of the display. In this study, we focus on the captured motive blur caused by the shutter speed of the camera. Generally, leveraging expensive high-speed camera equipment or using a de-blurring algorithm has been proposed to remove this type of blur. However high-speed cameras are too costly for the End-user to use, and de-blur algorithms have a problem that it takes quite a while to get remarkable results. Therefore we propose a method that uses a machine learning technique to obtain clear images even in low-end single RGB cameras with low frame rate.

Keywords: Machine learning · GAN (Generative Adversarial Network) · De-blurring · Motion blur

1 Introduction

We have so many impressive moments in our life such as traveling with friends or marrying someone in the family. With the smartphones, we have been able to take a variety of pictures for various purpose in various environments. However, there are cases in which important moments cannot be captured clearly due to restrictions on the shooting environment (illumination, camera shake, lens contamination, etc.). To solve the captured motion blur problem, we have to use expensive high-speed camera or variety of de-blur algorithms as described in the related work section. The cost of the

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 332–337, 2020.
https://doi.org/10.1007/978-981-13-9341-9_57

high-speed camera is too costly for End-users to use, and de-blur algorithms take a long time to obtain remarkable results. For these reasons, we propose a machine learning based de-blurring method which helps general users to get clear pictures without difficulty by using pre-learned weights.

In this paper, we use 'Image-to-Image Translation with Conditional Adversarial Networks ('Pix2Pix')' by Isola et al. [6]. Our main idea is to make a correspondence pair from the images taken by the ultra-high-speed camera and the corresponding blurred low framerate images to construct a learning data set. Then use the dataset to learn the difference between clear images and blurred images through Pix2Pix model. To train the Pix2Pix network, we used 240fps super fast image sequence included in the Need For Speed dataset [8]. The Need For Speed dataset consists of 100 high-speed videos (380k frames in total) taken in real life, and correspondingly, 30 fps videos artificially rendered through Adobe After effects are also included. Through this dataset, we want to perform self-supervised learning of the relationship between super-fast images and blurred images. And then do de-blurring action on images with motion blur captured by the End-user.

2 Related Work

There are three leading causes of blurring: caused by camera focus [3], caused by camera shake [1], and caused by the movement of the subject. If the image has DOF (depth-of-field) information, the blur caused by the defocused camera focus can be corrected to a clear image by using the optical characteristics which contained in DOF information. Blur due to camera shake can be divided into two parts: shake in one direction [2] and shake in multiple directions. Blur due to unidirectional shaking is analyzed by first and second derivative components in the estimated shake direction. After calculating the degree of shaking, a clear image is obtained [5]. This method can correct the little blur in a relatively short time because it is possible to perform the fast computation based on mathematics such as Fourier transform, which is not enough for zooming and motion compensation in directions not parallel to the subject [4]. If the camera moves in a direction that is not parallel to the movement of the subject path, or if the subject moves in a direction that is not parallel to the focal plane of the camera, and if the size of the subject changes as well as the direction of the subject (In this case when the unidirectional shake removing technique is used, the center portion of the image is relatively clear while the peripheral portion is not sufficiently corrected, so that the blurring still occurs.) is difficult to model the shake as a single variable. Therefore, the shake correction is performed by reflecting not only the motion of the shake in the horizontal direction but also the rotation, expansion, and contraction. However, in this case, the swing model is more complicated than the conventional unidirectional swing elimination technique, and the calculation speed is lowered so that it takes a relatively long time to obtain the result. To overcome this, researchers use learning methods such as deep learning instead of complex mathematical modeling based calculations.

3 Our Work

3.1 Redefine Dataset

Need For Speed dataset [8] contains over 100 240 fps high-speed images obtained from real life and synthetically rendered 30 fps blur images corresponding thereto. But since the blur intensity is too weak to utilize this dataset as it is, there was a problem that only little blur of the image was corrected. In the process of reducing the resolution of the original dataset with 256 * 256 pixels, which is the learning input resolution of the Pix2Pix model, some blur was lost and we failed to lead the training in the desired direction. To solve this problem, we modified the existing dataset and redefined it. First, the original high-speed image was reduced to 256 * 256, and after applying Adobe After Effect's Pixel Motion Blur to prevent blur loss due to resolution reduction. After that, we could obtain sufficient artificial blur image frames for training. We also learned about 41 scenarios out of a total of 100 videos. It contains three scenarios in which both the camera and the subject are moving, when the subject is moving with the camera stationary, and the camera is moving with the subject stationary. The newly redefined dataset includes blurred images that can occur in high and low illumination conditions and also contains indoor and outdoor environments. Too much learning data may interfere with learning, so we reduced the dataset to a total of 5439 frames. The testset consisted of 200 randomly extracted test data from all the training data, and the pair ground-truth images were extracted together to confirm the result after the estimation.

3.2 Learning

There are many different variants in the GAN, but in this study, we performed deblurring through 'Image-to-Image Translation with Conditional Adversarial Networks ('Pix2Pix') [6]'. Another study that performs image-to-image translation is Jun-Yan Zhu et al., 'Uncycling Image-to-Image Translation using Cycle-Consistent Adversarial Networks (CycleGAN) [7]'. This model learns common patterns between two different domains, is not suitable for our problem because we have to learn the difference between super fast images and synthetically rendered blur images.

Because of the loss of information when using the Encoder-Decoder structure, the authors of Pix2Pix used the U-Net architecture in the generator. Through this, we can concatenate image context of Contraction Path to Expanding Path to constitute Skip-connection structure, and as a result, image detail and image shape can be adequately maintained. Pix2Pix learns mapping function between one image domain to another image domain and verifies it through Discriminator. Interestingly, PatchGAN is applied to the Discriminator. Existing GANs determine the fact of genuine and fake by the whole image, but PatchGAN determines whether it is genuine and fake on a patch-by-patch basis. So it can send feedback to the generator from more detail. This study aims to learn the relationship between super-fast image and blurred image, and we can say this structure of Pix2Pix network is effective to solve our problem. Model training is performed through blur image sequence A (which is synthetically rendered by Adobe After Effect) and super fast image sequence B. The learning time increases linearly with the image length. The hyperparameters used in training are shown in the Table 1.

Table 1. Hyper parameters used in model training.

Hyperparameter	Value	Hyperparameter	Value
Aspect ratio	*1.0*	Beta 1, 2 for Adam	*0.5/default*
Batch size	*1*	NGF, NDF *	*64/64*
L1 weight	*100.0*	Learning Rate	*0.0002*
GAN weight	*1.0*	Epochs	*20/50/72/100*

*NGF = # of G filters in 1st CONV layer, NDF = # of D filters in 1st CONV layer

The computer used in training is Intel Core I9-9700 k, 32 GB of RAM and NVIDIA GTX 1080Ti (11 GB). Training takes about 7.5 min to perform 1Epoch, and it takes about 12.6 h to perform Max_Epoch 100. To compare the de-blurring performance according to Epochs, we performed inference through each testset at 20, 50, 72, and 100 epoch, respectively. The test result image is shown in the Fig. 1.

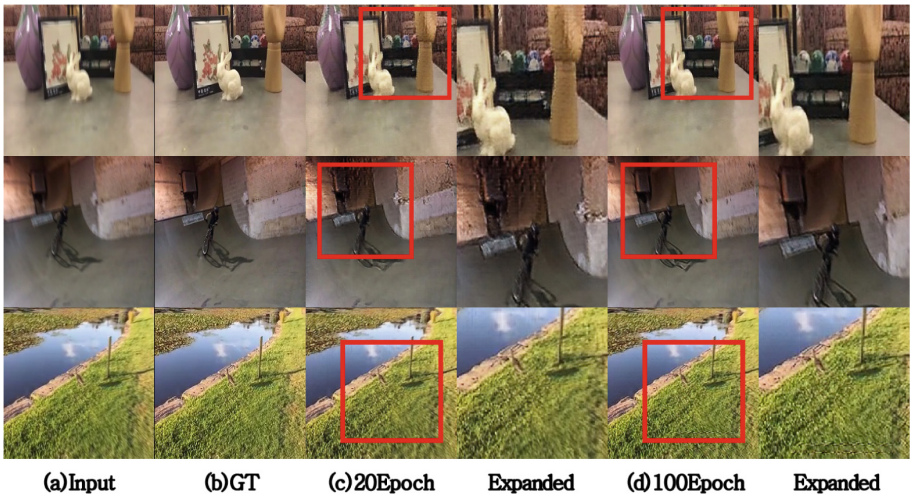


Fig. 1. These images show deblurring result from our method. (a) is a Test Input, (b) is Ground Truth Image, (c) and (d) shows the results of 20 and 100 epoch training, respectively.

We can confirm that the optimal number of epochs is 100 in the training process.

4 Conclusion

In this study, we proposed a new method to perform the deblurring by learning the relationship between high-speed image and blurred image using Pix2Pix which is one of self-supervised learning. Unlike conventional de-blurring methods that require complicated mathematical modeling processes for efficient blur elimination, using

machine learning such as GAN, the generators and discriminators compete with each other and repeat the process of learning the distribution of data. It allows us to solve the problem more concisely. However, if learning proceeds in an undesired direction, strange results can be obtained. And the limitations that can be considered as the cause of such problems can be found during the study. If we improve these limitations and extend our research, we're sure many people will be able to get a blur-free image easily without using expensive equipment.

4.1 Limitations

Time coherence and the problem of incomprehensible Image Context: At present, our approach cannot learn the temporal coherence of the image and the image context that exists in the image. If the image contains color gradation, not blur, but our model treats gradation as blur then generating strange results. If the subject is severely shaken and cannot distinguish the edges (the case that shape of the subject is broken due to severe blur), the model performs highly unstable estimation. (This can also happen when a residual image occurs due to a slow shutter speed). This problem leads to the generation of color artifacts, presumably due to the Mode Collapse problem and the lack of learning sample data.

4.2 Future Work

In future studies, we will apply FlowNet2 [9] method to the network structures to compensate for temporal artifacts and color artifacts caused by an incomplete understanding of Time Coherence and Image Context. And using that to perform accurate object segmentation and intermediate frame syntheses between deblurred images to turn blurred video into the ultra-high-speed video. The lack of datasets lead to the inaccuracy of deblurring. We will extend the data set to enable for deblurring in more cases. In general, taking a high-speed object with a low-end camera will cause the Jello Effect. By modifying the layers and hyperparameters of the Pix2Pix model used in this study, we expect that expanding to a robust Super-Slomo video generator that corrects not only the blur but also the Jello effect.

References

1. Holey, S.K., Warkar, K.V.: Enhancing video deblurring using efficient fourier aggregation-a review. Int. J. Eng. Sci. Res. Technol. **6**(12) (2017)
2. Takeda, H., Milanfar, P.: Removing motion blur with space–time processing. IEEE Trans. Image Process. **20**(10), 2990–3000 (2011)
3. Couzinié-Devy, F., Sun, J., Alahari, K., Ponce, J.: Learning to estimate and remove non-uniform image blur. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition (2013)
4. Cho, S., Lee, S.: Fast motion deblurring. ACM Trans. Graph. **28**(5), Article 145 (2009)
5. Choi, H.Y., Park, I.K.: Multi-view image deblurring for 3D shape reconstruction. J. Inst. Electron. Eng. Korea **49**(11), 47–55 (2012)

6. Isola, P., Zhu, J.-Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: Computer Vision and Pattern Recognition (2017)
7. Zhu, J.-Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: International Conference on Computer Vision (2017)
8. Kiani Galoogahi, H., Fagg, A., Huang, C., Ramanan, D., Lucey, S.: Need for speed: a benchmark for higher frame rate object tracking. arXiv preprint arXiv:1703.05884 (2017)
9. Ilg, E., Mayer, N., Saikia, T., Keuper, M., Dosovitskiy, A., Brox, T.: FlowNet 2.0: evolution of optical flow estimation with deep networks. In: Computer Vision and Pattern Recognition (2016)

Design and Implementation of an Enhanced Certified Document Archive System Based on PDF

Hyun Cheon Hwang[1], JiSu Park[2], Hyoung Guen Kim[1], and Jin Gon Shon[1(✉)]

[1] Department of Computer Science, Graduate School, Korea National Open University, Seoul, South Korea
{panty74, hgrikim, jgshon}@knou.ac.kr
[2] Division of General Studies, Kyonggi University, Suwon, South Korea
bluejisu@kyonggi.ac.kr

Abstract. An enterprise company produces a massive number of digital documents. Especially, a company which handles personal information such as insurance enterprise creates a personal digital document. This digital document needs to be certified and archived for a dispute between a company and a customer. And the enterprise needs to increase their archive storage because of a long-term archive. In this paper, we defined new PDF object keyword and proposed ECD-archive (Enhanced Certified Document archive) system which finds common resources from a certified PDF file and saves ECD-archive format to reduce storage space. Our experiment has shown ECD-archive system reduces entire storage space efficiently.

Keywords: Document archive · Certified document · ECD-archive system PDF · Digital signature document

1 Introduction

A certified document is a document which contains customer personalized contents or legal contract information with the enterprise. A certified document from an enterprise company such as financial enterprise to a customer has been becoming more critical parts in enterprise business management. The rate of legal depute between an enterprise and a customer has been increased, and the certified document such as digital signature PDF document is used widely in all customer communication due to a certified document has the same right with physically signed paper legally [1].

There has been a digital archive system such as reCAPTCHA for an ancient book or Web archive since the 1990s in public. And many enterprises have been used private archive system for a transaction log or a various document as well. A certified document archive system in the enterprise company has become one of the major systems in an enterprise because of regulation or dispute. However, the current document archive system consumes enormous storage space for archiving even though a certified document for each customer consists of both a small portion of personalized data and same

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 338–343, 2020.
https://doi.org/10.1007/978-981-13-9341-9_58

layout. In this paper, we proposed the enhanced certified document archive system which can reduce a storage space without any harm of document integrity while observing the digital archive standard.

2 Related Research

2.1 Certified Document

A certified document is a document which is be certified as true copies of the originals by an authority [2]. In an enterprise such as an insurance company, there are lots of electronic documents which are delivered from an enterprise to a customer. And these documents contain personal information and need to be assured the content came from the enterprise without any unauthorized modification. The electronic certified document based on PDF use digital signature technology to assure the document content. An enterprise can do sign on a document by using a private key and the digital signed document will be delivered with a certificate which has public key [1]. Korea government accept a digital document for using legal purpose instead of a physical paper document.

A certified document is created by document creation such as Adobe and digital signature solution such as HSM in an enterprise company.

2.2 Digital Archive System

An archive is an accumulation of historical records or the physical place they are located [3]. In the digital archive, there are several different domains such as digital document archive, web archive [4]. The CAGR of the data center storage is 31% during 2016–2021, and an archive system is also one of business application which uses a storage capacity [5]. Especially, an archive system uses huge storage capacity due to the characteristic of the system.

There are several recommended standard digital file formats such as TIFF, PDF [7] for document archive due to preventing digital dark age [6]. TIFF was used to be the most popular format, but PDF is more preferred due to it has text-searchable, independent to a device and digital signature inside of the file. Most of the case, the certified document system, and the archive system is separated for system stability. And the archive system must archive and retrieve this certified document without any defect to a digital signature.

2.3 PDF Document

PDF (Portable Document Format) is an OS and hardware independent format which is developed by Adobe [8]. It can be used in most of the system with the same layout, so many documents processing software supports PDF. PDF contains every data as an object and each object is consist of a tree structure. An object is one of many resource types such as a text, font, image, and each object has their parent object and child

```
5 0 obj
<< /Type /Exmple
    /Version 0.012
    key_n value_n >> stream end obj
```

Fig. 1. An object in PDF example

object [9]. Each object has dictionary entries to present the object properties, and more than one key-values pairs are written in double angle brackets (<<and>>) as shown in Fig. 1.

The advantage of the archive system with PDF is that PDF format is an independent document format for OS, devices, and the certificate PDF with a digital signature can guarantee the document integrity against unauthorized document modification. However, An archive system in an enterprise company needs a large storage system [10]. Enterprise can create, and archive one single bulk PDF document to prevent duplicated creation of the same resources such as a company logo by using batch processing for saving a storage space [11]. However, the document integrity with the digital certificate for each customer's document is not guaranteed by batch processing due to the bulk PDF document is not the same as the original document. In this paper, we propose ECD-archive (Enhanced certified document archive system) based on the proposed PDF object keyword, which can reduce storage space significantly without any harm for the digital signature on a certified document.

3 Design of ECD-Archive System

3.1 Extended Object Keyword in PDF

A document from an enterprise has personalized content area such as a customer name, address and common content area such as a company image, company address. ECD-archive detects the common content area from a document and extracts the content to reduce total storage space. We define new object keyword, ComRes to identify the object which as the shared content efficiently as shown in Table 1. Below the Fig. 2 shows how a text object structure can be shown with ComRes key.

Table 1. Key definition for shared resources

Key	Type	Value
ComRes	Integer	Shared resources: 1/else: 0

```
5 0 obj << /Type /XObject
           /ComRes 1
           /Subtype /Image >> stream end obj
```

Fig. 2. An object in PDF example with ComRes key-value

3.2 ECD-Archive File Format

Each certified document in an enterprise has the same shared resources such as a company logo and address. The certified document should have all resources in one single document to guarantee document integrity with a digital signature. However, it needs huge storage space to archive whole customers certified document in terms of a document archive system, and these documents need to be archived in the long-term period due to these documents could be used for proof against illegal dispute.

ECD-archive document splitter parses and extracts shared common resources from a document by using the extended pdf object keyword. And ECD-archive document splitter makes ECD-archive file after adding ECD-archive file header and description fields. ECD-archive file format is as shown in Tables 2 and 3.

Table 2. ECD-archive file format for main pdf document

Field		Byte	Value	Description	Occurs
ECDA file header indicator		10	%!ECDADOCU		1
Length of shared resource object description		8		Total description length	1
Shared resource object description	Offset	8		Repeated by the number of shared resource objects	n
	Length	8			
	Object ID	255	Unique ID		
	Checksum	32	MD5 checksum		
PDF page object		n	Binary		1

Table 3. ECD-archive file format for common resources

Field	Byte	Value	Description
ECDA file header indicator	10	%!ECDAOBJS	Only for common resources
Shared resource object ID	255	Unique ID	For all obj.
Shared resource object checksum	32	MD5 checksum	For all obj.
PDF page object	n	Binary	For all obj.

ECD-archive document splitter creates one ECDA main file and multiple ECDA resource files. ECDA main file has ECDA header and the pdf object which shared common resources are excluded. ECDA resource file has ECDA header and the shared common resources object. ECDA resource file can be created more than one by following the number of shared common resources in one certified document as shown in Fig. 3.

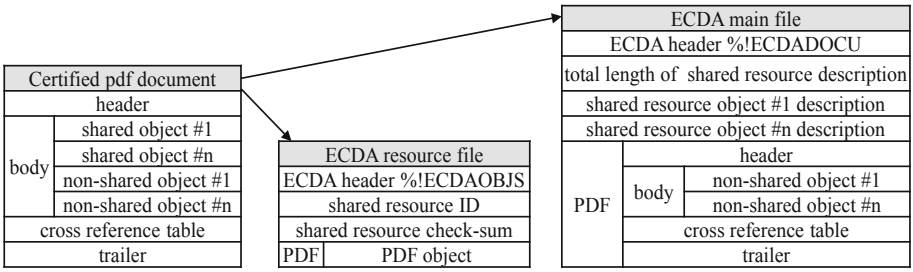


Fig. 3. Splitting a certificated document into ECD-archive files

3.3 Process of ECD-Archive System

All ECD-archive system stores all binary object data into DBMS for better data stability than storing into a file system. ECD-archive splitter skips storing the ECDA resource file in case the file is already stored in the DBMS by a previous archive process. Figure 4(a) shows the storing process of the certified document archive.

ECD-archive document composer recomposes a certified document from the DBMS. Recomposing process is (1) ECD-archive retrieve the ECDA main file (2) finds related ECDA resource files (3) verify checksum and object ID for recomposing integrity (4) merge these files into one certified document which is same with the original certified document as shown in Fig. 4(b).

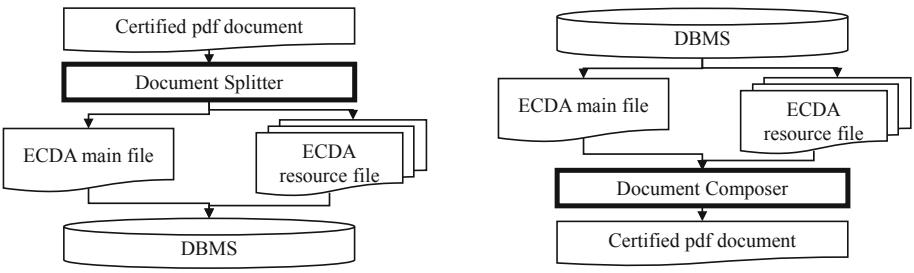


Fig. 4. Certified document process in ECD-archive

4 Experiment and Analysis

We verify the ECD-archive system described in this paper through this experiment. The test document properties are shown in Table 4. The total file size is about 28 times smaller than standard archive which is multiple PDF files because ECD-archive has only one shared resource file as shown in Table 5. The difference in overall file size is linearly larger along with the number of customers who have shared resources. Therefore, the ECD-archive system can reduce total storage space regarding the document archive system in an enterprise company.

Table 4. Experiment test document attributes

Shared resource	One image file (png, 2081 × 3508)
Total no. of pages	1 page
Total no. of customers	10,000 customers (1 file per each customer)

Table 5. Experiment result

Item	Normal archive	ECD-archive	Reduction ratio
Total file size	2,333 MB (10,000 files)	82 MB (10,001 files)	96.5%

5 Conclusion

In this paper, we researched and experimented ECD-archive system which extracts and stored shared common resources separately from a certified document for archiving and recomposes document to reduce storage space. As a result, efficiency is increased in terms of storage capacity. In conclusion, ECD-archive system is an efficient system for a certified document without any defect to document integrity. Especially when the number of a customer is enormous, and each customer uses the same document layout which has many same resources, the ECD-archive system shows better efficiency. In future research, we will research how ECD-archive system can reduce the overhead to process document-splitter and composer in terms of processing perspective.

References

1. Lim, Y.F.: Digital signature, certification authorities and the law. Murdoch Univ. Electron. J. Law **9**(3) (2002)
2. www.education.gov.au: Guidelines for the certification of documents. https://www.education.gov.au/guidelines-certification-documents. Accessed 10 Oct 2018
3. En.wikipedia.org: Archive. https://en.wikipedia.org/wiki/Archive. Accessed 20 Oct 2018
4. Kim, H.: A study on digital archiving and web archiving research analysis. Digit. Libr. **67**, 52–62 (2012)
5. Cisco: Cisco Global Cloud Index. https://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.pdf. Accessed 25 Oct 2018
6. Kuny, T.: The digital dark ages? Challenges in the preservation of electronic information. Int. Preserv. News **17**, 8–13 (1998)
7. www.lexisnexis.com: File formats for electronic document review. https://www.lexisnexis.com/applieddiscovery/lawlibrary/whitePapers/ADI_PDFTrumpsTIFF.pdf. Accessed 25 Oct 2018
8. www.adobe.com: PDF reference. http://www.adobe.com/devnet/pdf/pdf_reference.html. Accessed 25 Oct 2018
9. Adobe Systems Incorporated: PDF reference sixth edition (2006)
10. Bradshaw, P.L., et al.: Archive storage system design for long-term storage of massive amounts of data. IBM J. Res. Dev. **52**(4.5), 379–388 (2008)
11. Hwang, H.C., et al.: The design and implementation of an enhanced document archive system based on PDF. In: Advanced Multimedia and Ubiquitous Engineering, pp. 531–537. Springer (2018)

Comparison of Motion Similarity Using Image

Yunsang Jeong[1], Jinu Kim[2], Fajrul Norman Rashid[3], Lim Kok Yoong[3], Dongho Kim[1], and Jinho Park[1(✉)]

[1] Department of Media, Soongsil University, Seoul, Korea
kunka827l@gmail.com, c2alpha@gmail.com, cg@ssu.ac.kr
[2] Department of ICMC Convergence Technology, Soongsil University, Seoul, Korea
kjwl008@ssu.ac.kr
[3] Department of Creative Media, Multimedia University, Cyberjaya, Malaysia

Abstract. Recent motion recognition devices are widely used as inputs to analyze the interaction with motion and sports and entertainment content. However, the performance and efficiency of the conventional motion recognition apparatus are degraded despite the good quality information using a plurality of cameras and sensors. In this paper, we propose an algorithm that can be used as input means for posture analysis, posture supplementation, and interaction with contents in sport and entertainment by overcoming limitations of conventional motion recognition devices and using specific posture similarity judgment.

Keywords: Vision · Image processing · Motion comparison

1 Introduction

Recently, motion recognition devices are widely used as input for analyzing motion and interaction with content in sports and entertainment fields. However, existing motion recognition devices are composed of several cameras and sensors to recognize the space and the user, so there is not a small cost to use them and since it is used after fixing the device, it is difficult to use it effectively compared with the performance of the device.

In this paper, we propose an algorithm to overcome the above limitation by using only one RGB camera to recognize the user's motion and to compare the motion of a user with a predefined 3D model for the purpose of inducing a specific motion and determines similarity. The proposed algorithm is expected to be used as an input for interaction and assisting user's motion analysis and user's motion compensation in various sports and related entertainment contents using similarity.

2 Our Method

2.1 GMM-Based Background Subtraction

The background subtraction method based on GMM (Gaussian mixture model) [1–3] can be divided in to three steps: Gaussian distribution initial parameter setting,

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 344–349, 2020.
https://doi.org/10.1007/978-981-13-9341-9_59

foreground estimation, and GMM parameter update for foreground estimation of the next frame. But here, we do not update the GMM parameters since we assume that user need to keep certain actions for a while in order to recognize user's actions and that all situations happen indoors where the environment is almost unchanged. The background frame to be compared with the frame of the specific time is always the initial frame.

The GMM-based background subtraction algorithm distinguishes the background from the foreground using K Gaussian distributions for the pixels on the time axis in the mode pixel. The probability that the value of a particular pixel at time t has X_t can be written as

$$\mathrm{p}(X_t) = \sum_{i=1}^{K} \omega_{i,t} * \eta\left(X_t, \mu_{i,t}, \sum_{i,t}\right) \tag{1}$$

where K is the number of distributions, ω, μ, $\sum$(2) are the weight, average, covariance and normal distribution of the ith Gaussian component at time t.

$$\sum_{i,t} = \sigma_{i,t}^2 \mathrm{I} \tag{2}$$

And η(3) is a Gaussian probability density function.

$$\eta\left(X_t, \mu_{i,t}, \sum\right) = \frac{1}{(2\pi)^{n/2}|\sum|^{1/2}} e^{-\frac{1}{2}(X_t-\mu_t)^T \sum^{-1}(X_t-\mu_t)} \tag{3}$$

The Gaussian distributions are arranged according to the fitness value ω_k/σ_k [4]. The background value is relatively constant with the pixel value compared with the foreground, and the standard deviation is small and the weight value is large, so the fitness value is large. The first background distribution B(4) is a b-Gaussian distribution that is larger than a certain threshold value T, and others are regarded as foreground.

$$\mathrm{B} = \mathrm{argmin}_{\mathrm{b}}\left(\sum_{\mathrm{i}=1}^{\mathrm{b}} \omega_{i,t} > T\right) \tag{4}$$

2.2 Noise Reduction

We use only the initial background to estimate the foreground without updating the background. Therefore, minimizing the frame noise or the recognition and reduction of the change of the small environment is a way to express the foreground to be tracked more accurately.

We remove the noise by computing the image obtained from the camera with the mask generated by using the two-dimensional Gaussian function value before performing the background subtraction. Since the two-dimensional Gaussian function can be divided into a product of a one-dimensional function, it is possible to implement a fast Gaussian filter by appropriately using it. If a function for creating a two-dimensional mask is divided into the directions of each axis, a one-dimensional mask operation can be performed to generate a result image.

We use erosion and dilation which are the basis of morphology operations [5], to show strong background subtraction performance when in the case of noise and minute environmental changes. First, the erosion is performed using the binary mask obtained through the background subtraction and the predefined structuring element. In this case, if structuring element completely falls within binary mask, the pixel value below structuring element's fixed point is set to 1. Here, the small noise that remains in the binary mask is disappear. Then, the dilation operation is performed using the eroded binary mask and the predefined structuring element. At this time, if structuring element overlaps with eroded binary mask, the pixel values at the position of structuring element set to 1.

2.3 Extracting Contour

The contour is extracted by combining the boundaries of consecutive points with the same color or intensity in the noise-canceled binary mask. Since we need only the human outline in the estimated foreground binary mask, we use the suzuki85 algorithm proposed by Suzuki et al. [6]. To collect only the components at the top of the hierarchical structure of the contours, and extract the closed curve with the largest area among them. The area of the closed curve is obtained using Green's formula (Fig. 1).

Fig. 1. The process of extracting contour, (a) background, (b) frames of a specific time, (c) the contour of foreground mask obtained in (b)

2.4 Motion Comparison

We have created an algorithm that detects the similarity of motion by using two images that have only motion left by deleting the background. First, the preprocessing process of converting the extracted motion image into black and white through binarization was performed. When comparing motion, there are two ways of preprocessing both the image extracted using the target motion and the image extracted from the motion of the user in real time and the preprocessing only the motion of the user. In the case of the former method, it is possible to extract faster and more accurate results when comparing operations using only 0 and 1 operations. In the latter case, weights can be

imposed for finer operations. Two preprocessed images will have a color value of 0 or 1 for all pixels. This feature makes it very easy to extract the similarity of motion between two images.

$$R = T_c * C_c. \tag{5}$$

First, the color value Tc of the processed image T and the motion of the user are extracted using the target motion as in Eq. (1), and the result is multiplied by the color value Cc of the processed image C and stored in R. The two preprocessed images are all 0 or 1, so all the color values are 0 or 1.

$$S = \sum R_n. \tag{6}$$

Next, the R values extracted through the Eq. (1) are added as shown in Eq. (2). S, calculated by adding all R, indicates how similar the two images are. If the S value thus extracted is compared with the sum of the operation color values of the original image T, the similarity of the two operations can be finally extracted.

```
program CompareSimilarity (Output)
  Color[] originColor = Origin Image Color;// Cc
  Color[] targetColor = Target Image Color;// Tc

  float[] resultColor = Size(originColor.Length);

  float originRate = 0;
  float resultRate = 0;

  for(int I = 0 ; I < originColor.Length; I ++){

      originRate += originColor;
      resultColor = originColor + targetColor;

      resultRate += resultColor;

}
   return resultRate / originRate * 100.f;
```

2.5 Verification of Motion Comparison Algorithm in Full Frame

We have used a motion comparison algorithm to verify that it can be performed in real-time as well as in a single image. The time was measured using motion comparison algorithm for each frame. We measured the time through several experiments and verified that the algorithms can be performed 120 to 140 times per second (Fig. 2).

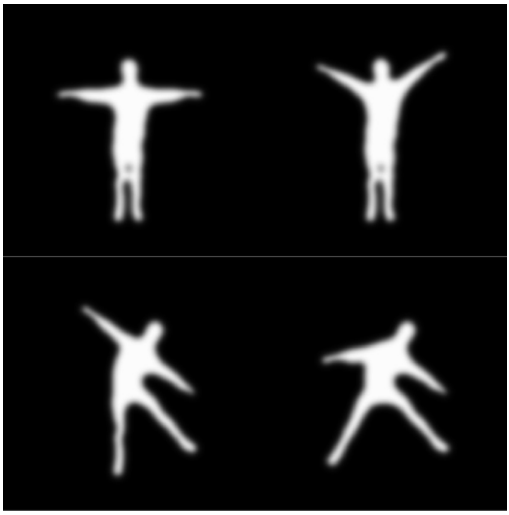


Fig. 2. Experimental images are in order Origin, Motion Image 0, Motion Image 1, Motion Image 2

Table 1. Verification of operation comparison algorithm

Camera image	Similarity	Time
Motion image 0	25.86	9 ms
Motion image 1	46.55	1 ms
Motion image 2	84.48	7 ms
Origin	100.0	1 ms

Table 1 shows the execution time when the comparison algorithm is performed using the image input from the camera and the target image of the motion. The execution time of each image did not exceed 12 ms and showed an average of 5 ms.

3 Result and Limitation

If we comprehensively use the method suggested by us, we expect to know how accurate the operation is being performed through the comparison of operations. In addition, it can process video and camera data in real time beyond a single image and can be utilized in various contents. It can play a sufficient role in helping the movement of yoga and pilates that require accuracy of movement. But the way we present does not pursue high accuracy. Although it can inform and guide the action, there is a disadvantage that it does not know whether the action is the same. We expect to be able to expand to more accurate research through machine learning or using various methods.

References

1. KaewTrakulPong, P., Bowden, R.: An improved adaptive background mixture model for real-time tracking with shadow detection. In: Remagnino, P., Jones, G.A., Paragios, N., Regazzoni, C.S. (eds.) Video-Based Surveillance Systems, pp. 135–144. Springer, Boston (2002)
2. Zivkovic, Z.: Improved adaptive Gaussian mixture model for background subtraction. In: Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004, vol. 2, pp. 28–31 (2004)
3. Zivkovic, Z., Heijden, F.V.: Efficient adaptive density estimation per image pixel for the task of background subtraction. Pattern Recogn. Lett. **27**, 773–780 (2006)
4. Stauffer, C., Grimson, W.E.L.: Adaptive background mixture models for real-time tracking. In: Proceedings of the IEEE Conference on CVPR, pp. 246–252 (1999)
5. Morphology. http://homepages.inf.ed.ac.uk/rbf/HIPR2/morops.htm
6. Suzuki, S., Abe, K.: Topological structural analysis of digitized binary images by border following. Comput. Vis. Graph. Image Process. **30**, 32–46 (1985)

OAuth-Based Access Control Scheme Against Replay Attacks in IoT Environment

Dae-Hwi Lee[1] and Im-Yeong Lee[2(✉)]

[1] Department of Computer Science Engineering, Soonchunhyang University, Asan-si, Korea
leedh527@sch.ac.kr
[2] Department of Computer Software Engineering, Soonchunhyang University, Asan-si, Korea
imylee@sch.ac.kr

Abstract. The OAuth protocol is a standard that authorizes not only the IoT environment but also the client services connected to the Internet to access the resource of the user stored in the resource server without sharing the account information of the resource server. Each IoT service uses OAuth to reduce the exposure of user privacy and provides scalability and flexibility because it can take advantage of any standard-compliant service. However, in recent OAuth 2.0 protocols, security vulnerabilities such as replay attacks and CSRF are already announced. To prevent this, a variety of OAuth-based access control protocols are being researched. In this paper, we design an OAuth-based access control protocol that is safe for replay attacks in IoT environment, so that IoT service can securely access the user resource of external resource server.

Keywords: IoT security · OAuth · Access control · User authorization

1 Introduction

Recently, the Internet of Things (IoT) environment continues to develop as it becomes the center of the fourth industrial revolution. As a result, various IoT services have been launched to make our life easier. In addition to IoT, cloud, big data, and mobile technologies are attracting interest and research as technology that is the core of the fourth industrial revolution. Smart home, healthcare, smart factory, and smart city service converge each technology [1]. These services are operated by connecting various sensor devices to the Internet and connecting sensor devices, communication technologies and interfaces. Major manufacturers in various countries are launching various types of IoT devices. OneM2M and IETF are working on standardization of IoT communication technology so that these various IoT devices can communicate smoothly.

However, in IoT environment, security should be considered as much as various devices communicate with various communication technologies. When various devices of different kinds are communicating, security vulnerabilities of each device are gathered, and new security vulnerabilities may occur. In addition, privacy problems may occur in the process of collecting data by sensor devices [2]. Especially privacy

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 350–355, 2020.
https://doi.org/10.1007/978-981-13-9341-9_60

issues are very important in IoT. For example, smart home service data is sensitive data that is most closely related to the daily life of a smart home user. This data can expose the user to monetary damage and can lead to personal injury. Furthermore, as it can identify power usage and cycle, gas valve lockout, and whether a user is in a healthcare device or home. have. Therefore, the issue of privacy is an important part of IoT.

In this paper, we propose a secure access control scheme for replay attacks in IoT environment where each service is interlocked by using OAuth as a tool for solving privacy exposure. OAuth, recently supported by many services, is a proposed standard in 2012 and authorizes the IoT service connected to the Internet to access the user's resources stored in the external resource server without sharing the user's information. This can solve the privacy problem by not leaving the user's information in the IoT service [3, 4].

2 OAuth (Open Authorization)

OAuth is a defined standard that can authorize a service to access user resources stored in an external resource server without sharing user information. Using OAuth, the client service authenticates the user through an external authorization server and then grants the client service access to the resource server. This makes it possible to integrate and use various client services without requiring a separate authentication process from the client service. OAuth was first introduced in 2010 by the IETF OAuth Working Group as OAuth 1.0 protocol standard RFC5849 [3]. Recently, a standard RFC6749 for OAuth 2.0 framework has been announced [4], and standardization work is progressing according to the communication scheme between each communication object [5].

[6] is the RFC 6819 standard that addresses security threats and considerations in the proposed RFC 6749 standard in 2012. RFC6749 introduces several security vulnerabilities in RFC6819, and typically includes replay attacks and CSRF (Cross Site Request Forgery) attacks.

3 Security Requirements

Access control: The client gets access to the resources of the resource server using the token issued by the authorization server through the user. Only a client communicating with an authorization server via a legitimate user must acquire the privilege and securely receive the token. In this process, replay attacks and CSRF attacks should be considered.

Prevention of replay attacks: During the process of receive the token at client, the user's information and authentication code must not be retransmitted. If an unauthorized attacker intercepts and reuses token, it should not be able to get a token.

Prevention of CSRF attacks: To prevent CSRF attack, it is necessary to check whether the user who generates the request message and the user who receives the token are identical. When a user sends a request message to a client while using a

service, a CSRF attack should be prevented by checking whether the user who issued the token is using the same parameter as the state.

4 Proposed Scheme

The OAuth-based access control scheme for the IoT environment proposed in this paper is divided into two phases. Figure 1 is a scenario of the proposed scheme. The client is pre-registered in the authorization server and shares the symmetric key.

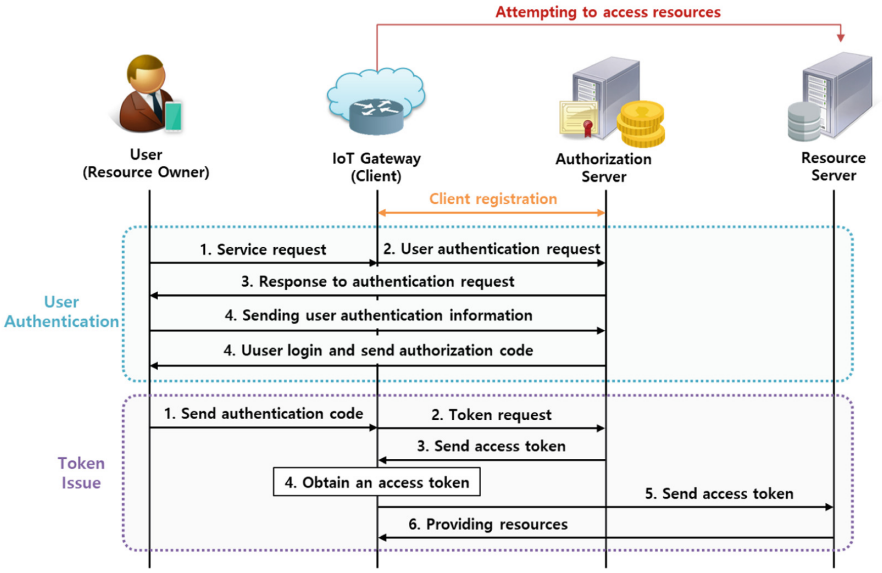


Fig. 1. Scenario of proposed scheme

4.1 System Parameters

- $*$: Participant object
  (U: User, C: Client, AS: Authorization Server, RS: Resource Server)
- $*_{ID}$: The identifier of the participating object
- $h(\cdot)$: One-way hash function
- r: The random number used for the current session created by the client
- $K_{AS,C}$: The symmetric key shared by the authorization server and the client
- $Session_{ID}$: The current session identifier issued by the authorization server
- TS: Time stamp
- LT: Lifecycle of the token and the resource content requested to the resource server

4.2 User Authentication and Login Phase

Step 1. The user transmits a hello message nonce to the client through the PC or mobile device to receive the service of the client using user's resource of the resource server.

Step 2. The client receiving the Hello message registers the client and sends the following request message to the authorized server sharing the symmetric key $K_{AS,C}$. The r is an arbitrary number to be used in the current session, and the authorization server is an authorization server managed by the resource server in which the resource of the user is stored.

$$Request\left(nonce+1 \parallel E_{K_{AS,C}}(r \parallel U_{ID})\right)$$

Step 3. The client sends the message $Hello(nonce+1 \parallel URI_{AS} \parallel r)$ containing the r and the URI of the authorization server in response to the hello message to the user.

Step 4. The authorization server checks the request message sent by the client, identifies the user's identity, and sends the identity AS_{ID} of the authorization server to the user to request authentication information of the user.

Step 5. After confirming the identity of the authorized server, the user redirects through the URI_{AS} and transmits the message $\mathrm{U_{ID}} \parallel h(U_{PW} \parallel r \parallel nonce+1)$ generated by the user's ID and password.

Step 6. The authorization server checks the user's ID and password and constructs a secure channel through HTTP-TLS or CoAP-DTLS handshake. The authorization server sends the following message over the secure channel. $Session_{ID}$ is a session identifier for authentication of the authorization server in the current session, $AUTH$ is the authentication code for issuing the token of the client, and TS is the current time stamp.

4.3 Token Issue Phase

Step 1. The user sends the session identifier $Session_{ID}$, the authentication code $AUTH$, and the time stamp TS to the client as well as the nonce value initially generated.

Step 2. The client sends the following message including the session identifier $Session_{ID}$, authentication code $AUTH$, and time stamp TS received from the user to the authorization server.

$$\mathrm{E_{K_{AS,C}}}(Session_{ID} \parallel AUTH \parallel TS)$$

Step 3. The authorization server decrypts the received message and checks whether the authentication code matches the session identifier issued by the authorization server. If the verification is successful, the next token is encrypted and sent to the client.

$$\mathrm{E_{K_{AS,C}}}(E_{PR_{AS}}(h(Session_{ID} \parallel C_{ID}) \parallel Session_{ID} \parallel C_{ID} \parallel TS \parallel LT))$$

Step 4. The client receiving the encrypted token decrypts the token and transmits it to the resource server. After confirming the token using the public key of the authorized server, the resource server checks the identifier C_{ID} of the client and forms a secure

channel with the client. The resource server then provides the user's resources available to the client for the access rights specified in the *LT*. Future *LT* and *TS* will be used to manage the validity period and privileges of the token session.

5 Analysis of Proposed Scheme

Access control: The client receives a valid token through the authorization server and the user. The client that has issued the token obtains the right to access the resource server. This makes it possible to securely obtain the resources of the user available at the client. The token can be issued only by the user's authentication information using the symmetric key $K_{AS,C}$ to the client registered in the authentication server.

Prevention of replay attacks: An arbitrary number r of the request message transmitted by the client to issue a token to the authorized server is a value that can be used only in the current session. Since the client transmits the r to the user and redirects the user to the authorization server, the client transmits the hash value for r and the nonce together with the authentication information of the user. Therefore, the authorization server authenticates only the user who transmitted r in the current session. This value cannot be used in the next session even if the attacker seizes it and retransmits it in the middle.

Prevention of CSRF attacks: The token issued by the authorization server contains the $Sessoin_{ID}$, which is the identifier for the session in which the current token is used, and the C_{ID} of the client's identifier and contains the signature value encrypted with the private key of the authorization server. Therefore, the contents included in the token proves that the authorization server has issued and can be informed that it has been issued to the specific client C_{ID}. Then, when the client provides the service to the user, the user can prevent the CSRF attack by generating a request using the $Session_{ID}$ (Table 1).

Table 1. Comparison of proposed scheme with existing schemes

	Kothmayr scheme [7]	Sciancalepore scheme [8]	Proposed scheme
OAuth structure	Authorization code	Implicit	Implicit
Prevent replay attacks	△ Depend on HTTPS	○ Use HTTPS & HMAC	○ Use HTTPS & nonce
Prevent CSRF attacks	X	X	○

6 Conclusions

In this proposed scheme, we propose OAuth - based access control scheme which is safe for replay attacks in IoT environment. It satisfies the structure of the OAuth standard by applying the proposed scheme scenario shown in Fig. 1 and is designed to be safe for replay attacks and CSRF attacks in the process of transmitting each message. In the proposed IoT environment, the number of communications is relatively high, and the number of cryptographic algorithms used is higher than that of OAuth. However, it is considered safe for replay attack and CSRF attack, Obtain access privileges.

OAuth has been undergoing a lot of research recently due to the development of IoT environment, and it is a technology applied by various services, and continuous research should be conducted to meet more diverse environments. In the future, it will be necessary to study the direction to reduce the number of operations and to provide more efficiency in the IoT environment. We will compare and analyze the implementation through actual implementation. IoT service is closely related to our life, so if security is weak, it can cause a big accident. Therefore, many researches on access control techniques for IoT environment should be made.

Acknowledgement. This research was supported by the MSIT (Ministry of Science, ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2018-2015-0-00403) supervised by the IITP (Institute for Information & communications Technology Promotion).

References

1. Karaköse, M., Yetis H.: A cyberphysical system based mass-customization approach with integration of Industry 4.0 and smart city. Wirel. Commun. Mob. Comput. (2017)
2. Ziegeldorf, J.H., Morchon, O.G., Wehrle, K.: Privacy in the Internet of Things: threats and challenges. Secur. Commun. Netw. **7**(12), 2728–2742 (2014)
3. Hammer-Lahav, E.: The OAuth 1.0 Protocol. IETF. RFC5849 (2010)
4. Hart, D.: The OAuth 2.0 Authorization Framework. IETF. RFC6749 (2012)
5. Tschofenig, H.: The OAuth 2.0 Internet of Things (IoT) Client Credentials Grant. IETF Internet Draft. draft-tschofenig-ace-oauth-iot-00.txt (2014)
6. Lodderstedt, T., McGloin, M., Hunt, P.: OAuth 2.0 Threat Model and Security Considerations. IETF, RFC6819 (2013)
7. Kothmayr, T., Schmitt, C., Hu, W., Brünig, M., Carle, G.: DTLS based security and two-way authentication for the Internet of Things. Ad Hoc Netw. **11**(8), 2710–2723 (2013)
8. Sciancalepore, S., Piro, G., Caldarola, D., Boggia, G., Bianchi, G.: OAuth-IoT: an access control framework for the Internet of Things based on open standards. In: 2017 IEEE Symposium on Computers and Communications (ISCC), pp. 676–681 (2017)

A Study on Secure Data Access Scheme Based on CP-ABE in Cloud Environments

Yong-Woon Hwang[1] and Im-Yeong Lee[2(✉)]

[1] Department of Computer Science Engineering, Soonchunhyang University, Asan, South Korea
hyw0l23@sch.ac.kr

[2] Department of Computer Software Engineering, Soonchunhyang University, Asan, South Korea
imylee@sch.ac.kr

Abstract. Recently, with the development of cloud computing, it has become possible to store and share data in cloud environments; however, there are various security threats in these environments. Attacker can leak stored data and, above all, service providers cannot be trusted completely. Therefore, security technologies guaranteeing that data remains secure when stored in cloud environments are required. To this end, researchers are studying techniques based on CP-ABE. However, vulnerabilities to various security threats remain and some schemes are inefficient. In this paper, we apply CP-ABE to address some of the security threats associated with cloud environments, by ensuring that only authorized users can access data stored in the cloud. In addition, we propose an access control system that, in comparison to existing schemes, is more computationally efficient when attributes are revoked.

Keywords: Cloud · Access control · Attribute-based encryption · Revocation

1 Introduction

Recent developments in cloud computing, such as the products offered by Amazon and Google, allow people to store data on cloud platforms or share them with others as needed, rather than having to purchase storage at cost. For this reason, cloud storage is widely used in real life environments, including by medical, company, and public institutions. However, there are various security threats present in cloud environments. In particular, service providers cannot be trusted completely, and malicious users can leak or lose data. As a result, we need to ensure the security of the data owners' data in the cloud environment and develop technologies enabling users to securely access cloud storage. Attribute-based encryption (ABE) is one of the most suitable cryptographic technologies for implementing access control in cloud environments, and ciphertext-policy ABE (CP-ABE) is widely used [2]. Research on access control using CP-ABE in cloud environments has shown continual progression. However, some schemes are vulnerable to spoofing attacks, while others are inefficient because all keys and ciphertexts must be updated when a user is revoked.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 356–362, 2020.
https://doi.org/10.1007/978-981-13-9341-9_61

In this paper, we propose an access control technology that allows authorized users to access cloud storage facilities safely via the CP-ABE scheme.

In addition, we improve the efficiency of the scheme by ensuring that existing users' keys and ciphertexts are not affected when attributes are revoked.

2 Related Work

In this section, we discuss papers in which CP-ABE has been applied within an ABE cloud environment.

2.1 Attribute Encryption

Attribute-based cryptography is used to perform encryption/decryption based on a set of attributes (such as company affiliation or job), which are pieces of information about various stored objects; the access structure (AS) is based on this information. In the CP-ABE scheme, when a sender specifies the AS based on a recipient's attributes, and sends it to the recipient together with the encrypted data, the receiver decrypts the ciphertext based on its attribute set [1].

2.2 The Proposed CP-ABE Scheme in the Cloud Environment

As shown in Fig. 1, many papers have proposed data access techniques based on the CP-ABE scheme in the cloud environment [2–6], but some of the schemes developed are vulnerable to various security threats, while others lack efficiency. In particular, the schemes of Sekhar [2] and Zhu [3] are vulnerable to camouflage attacks. These can occur when a user can access data via another user's attributes. In response to this, some studies have applied attribute revocation when a user withdraws, such as Xu [4], Yang [5], and Ramesh [6]. However, the schemes of Xu and Yang are ineffective with respect to updating other users' keys and existing stored ciphertexts upon user withdrawal. Also, the scheme of Ramesh is inefficient because it continuously updates the public attribute keys of other users within the attribute authority (AA), which manages the withdrawal of attributes.

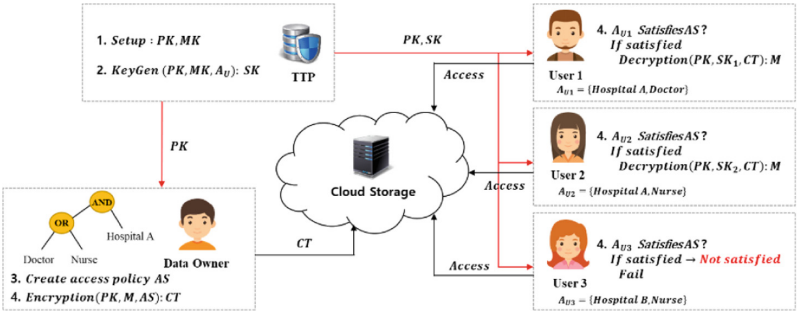


Fig. 1. The CP-ABE model in a cloud environment.

3 Proposed Schemes

In this section, we propose a data access control scheme based on the CP-ABE technique in a cloud environment, which is summarized in Fig. 2. The proposed scheme cannot be used to access an attribute in cloud storage after it has been withdrawn by the user using the attribute revocation function. Within the existing cloud environment, the CP-ABE approach allows users to access cloud storage facilities directly, while we add access control (AC) in our proposed scheme. This manages users' access to cloud storage facilities. Once a user is satisfied with the ciphertext access policy structure specifications, with respect to their own attributes and those of the data owner in the AC at the time of access, a key, denoted ak, is generated and sent to the user so that they can decrypt the ciphertext. Therefore, it is possible to increase the efficiency of the calculation by reducing the computational burden associated with conventional data decryption. The protocol operation pairing operation of the proposed scheme is used and developed based on BSW scheme [1].

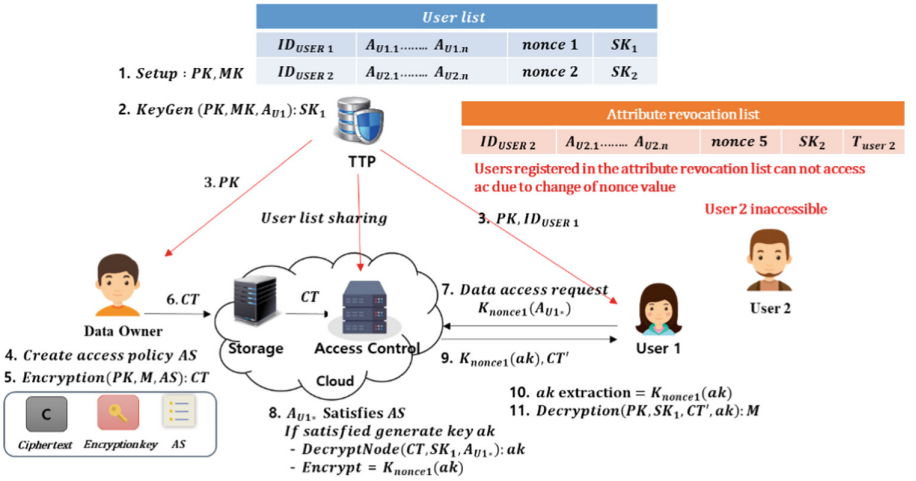


Fig. 2. Proposed data access control scheme

3.1 System Parameters

The system parameters used in the proposed scheme are as follows.

- TTP: manages user attributes with three trusted organizations
- Storage: servers that manage data
- Access Control: user access control management
- pk, mk, sk: Public parameter, master key, user security key (decryption key)
- A_U, S: user attribute data, a set of attribute data
- AS: structure of the ciphertext policy
- ak: decryption key generated when the user attribute matches the AS
- CT, CT': encrypted data, attribute matching verified CT

3.2 Proposed Scheme Scenario Phase

Step 1. When a user sends a message to the TTP to request registration, the TTP registers the information about the user in the user list and generates a public parameter and a master key.

- Setup(1^k): Generate public parameter $pk = \{G_0, \mathrm{g}, \mathrm{e}(\mathrm{g}, \mathrm{g})^a\}$ and master key $mk = \{B, g^a\}$

Step 2. An sk is generated based on the user attribute data, an mk, and a pk. It then sends the pk to the data owner and the user, ID_{USER}, registered on the user list via a secure channel.

- KeyGen(pk, mk, A_U): Generate secret key $sk = (D = g^{a+r}, \forall j \in A_{U*} : D_j = g^r H(j)^{r_j}, D_j^{'} = g^{r_j})$ through set of attributes A_U and mk. r is a random number value equal to the nonce value.

Step 3. The data owner specifies the AS based on the attributes of the user, who can access the data, m, to be shared. These data are then encrypted with public parameters and an AS (pk, AS), and are then transmitted to cloud storage

- Encrypt(pk, AS, m): Encrypt data m with pk and access structure AS, random number s
- $\mathrm{CT} = \left(AS, \mathrm{C}^{'} = \mathrm{Me}(\mathrm{g}, \mathrm{g})^{as}, \mathrm{C}^{*} = \mathrm{g}^{s}, \forall y \in \mathrm{AS} : C_y = g^{q_y(0)}, C_y^{'} = H(att(y))^{q_y(0)}\right)$

Step 4. The user requesting access to the encrypted data is encrypted with the nonce value received from the TTP and transmits the encrypted attribute set to the AC.

- $k_{nonce}(A_{U*})$: user requests access to AC by encrypting its attribute with nonce value

Step 5. The AC decrypts the $k_{nonce}(A_{U*})$ received from the user and compares the AS, as specified in the user's attribute set and ciphertext, to generate the key ak, which is necessary for decoding the data. It then encrypts ak with the user's nonce value and sends it to the user, along with the CT'.

- A_{U*} *Satisfies AS*? (*If satisfied*, *generate data decryption key* ak)
- DecryptNode $(CT, A_{U*}, r) : ak = \mathrm{e}(\mathrm{g}, \mathrm{g})^{rq_{root}(0)} = \mathrm{e}(\mathrm{g}, \mathrm{g})^{rs}$
- $k_{nonce}(ak), CT'$: sent to the user who satisfied the attribute

Step 6. The user decodes the $k_{nonce}(ak)$ received from the AC and thus acquires the ak. Hence, the cryptosystem CT decrypts message m using the private keys sk and ak, and the public parameter pk.

- Decrypt (pk, sk, CT', ak): the user decrypts the CT with the sk and ak
- $\mathrm{M} = \mathrm{C}^{'}/(\mathrm{e}(\mathrm{C}^{*}, \mathrm{D})/ak)$

Step 7. When a user withdraws, the TTP registers the user's information in the attribute revocation list, changes the nonce value, and shares it with the AC. Since the AC is not aware of that user's nonce value, upon making an access request based on the withdrawn user's information $(nonce, A_{U*})$, an attacker is unable to decode the received

$k_{nonce}(A_{U*})$. The AC then blocks the unsubscribed user from accessing the system by sending them an access failure message (Table 1).

Table 1. Comparison between existing CP-ABE scheme used in cloud environment.

Scheme items	Sekhar [2]	Zh [3]	Xu [4]	Yang [5]	Ramesh [6]	Proposed scheme
User collusion	Safe	Safe	Safe	Safe	Safe	Safe
Camouflage attack	Unsafe	Unsafe	Unsafe	Safe	Safe	Safe
User access control	Unsafe	Unsafe	Safe	Safe	Safe	Safe
User revocation	Not considered	Not considered	Satisfied	Satisfied	Satisfied	Satisfied
Attribute revocation	Not considered	Not considered	Not required	Satisfied	Not required	Satisfied
Key, ciphertext updates	Not considered	Not considered	Update all keys and ciphertext	Update all keys and ciphertext	Update all keys	Not required

4 Security Analysis of Proposed Scheme

- User collusion: The proposed scheme generates a secret key based on attributes and random values, so it cannot derive the value of *sk* from the user attributes alone. Also, the nonce value is known only to the user, and we use an AC as a symmetric key to request data access. The key *ak* is also used to decode data. Thus, even if *sk* is estimated via collisions between users, we cannot decode m.
- Camouflage attack: In the proposed scheme, the TTP removes the user ID and attribute from the user list when the user withdraws, and shares the update with the AC. Communication between the user and the AC is performed using the user's nonce value, which is registered in the user list as a symmetric key. Therefore, the AC cannot decrypt the access request message even if the withdrawn user accesses the AC with another user's attribute. Therefore, it is secure against camouflage attacks, where another person's attributes are used.
- Control access by unauthorized users: In the proposed scheme, the user is initially registered in the TTP to access the data, which also requests access to the AC so that the encrypted stored data can be decrypted. The AC compares the attribute of the user with the *AS* attribute specified in the encrypted data, and transmits the ciphertext to the matching user. Therefore, it is possible to prevent access by unauthorized users.

The proposed scheme has the following advantages compared with the existing CP-ABE scheme. It is safe against the camouflage attacks that can arise when applying the schemes of Sekhar [2] and Zhu [3]. It is inefficient to update all of the other users' keys

and existing stored ciphertexts, as occurs when a user withdraws from the schemes of Xu [4], Yang [5], or Ramesh [6]. By adding the attribute revocation list and AC, we have improved the efficiency of the proposed scheme because the other users' keys and ciphertext operations are not affected when a user withdraws. Finally, in the existing scheme, the user performs the attribute-matching operation directly on the specified *AS* when downloading and decrypting the cipher text. The proposed scheme can reduce the burden on the user's decoding computation because the AC performs part of the decryption operation.

5 Conclusions

In this paper, we propose a data access control scheme based on CP-ABE to solve security threats associated with cloud data storage. The proposed scheme is not vulnerable to the same spoofing attacks that affect other schemes, wherein other users' attributes can be employed to access data, because we use a nonce value, provided by the TTP, as a symmetric key for accessing the cloud storage facilities. Also, the AC grants users access to cloud storage based on the list of users managed by the TTP, which is updated every time a user is added or withdrawn. Hence, other users' keys and ciphertext updates are not affected by the addition or withdrawal of users. Therefore, in contrast to existing schemes, the attributes can be revoked efficiently.

In future studies, we will focus on improving the efficiency of our scheme by comparing its computations to those of existing schemes.

Acknowledgments. This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No. 2017-0-00156, The Development of a Secure Framework and Evaluation Method for Blockchain) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2016R1D1A1B03935917).

References

1. Bethencourt, J., Sahai, A., Waters, B.: Ciphertext-policy attribute-based encryption. In: IEEE Symposium on Security and Privacy, SP 2007, pp. 321–324 (2007)
2. Sekhar, B.R., Kumar, B.S., Reddy, L.S., PoornaChandar, V.: CP-ABE based encryption for secured cloud storage access. Int. J. Sci. Eng. Res. **3**(9), 1–5 (2012)
3. Zhu, S., Yang, X.: Protecting data in cloud environment with attribute-based encryption. Int. J. Grid Util. Comput. **6**(2), 91–97 (2015)
4. Xu, Z., Martin, K.M.: Dynamic user revocation and key refreshing for attribute-based encryption in cloud storage. In: 2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications, pp. 844–849. IEEE (2012)
5. Yang, K., Jia, X.: Attribute-based fine-grained access control with efficient revocation in cloud storage systems. In: Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security, pp. 523–528. ACM (2013)

6. Ramesh, D., Priya, R.: Multi-authority scheme based CP-ABE with attribute revocation for cloud data storage. In: 2016 International Conference on Microelectronics, Computing and Communications (MicroCom), pp. 1–4. IEEE (2016)
7. Stallings, W.: Cryptography and Network Security: Principles and Practice, Global edn. Pearson Education Limited, London (2016)

An Optimal Time Allocation to Maximize Total Average Transmission Data of Two-Way Relay System

Taehoon Kwon(✉)

Data Analysis Platform Center, Korea Institute of Science and Technology Information, 66, Hoegi-ro, Dongdaemun-gu, Seoul, Korea
kth78@kisti.re.kr

Abstract. In this paper, we proposed an optimal time allocation to maximize the total average transmission data of two-way relay system. First, we derived the average transmission data rate of two-way relay system by the probability of successful transmission and the target data rate. Next, based on the derived the average transmission data rate of two-way relay system, we proposed an optimal time allocation of each link as a closed form. The simulation result verifies that the proposed time allocation is optimal.

Keywords: Time allocation · Two-way relay · Optimization

1 Introduction

The relay system has recently gained much attentions related to the small cell technology, which is proposed as one of the solutions to support 5G requirements. In relay system, a relay is installed and operated in radio shadow area where it is difficult for a signal to reach such as interiors, undergrounds [1–3]. However, in a classical relay system where one node uses one relay exclusively, an extra time is required for relay to receive and transmit signals. It decreases a resource efficiency [3].

In order to prevent the resource waste, two-way relaying system as known bi-directional relaying, is proposed [4–7]. In two-way relaying, one relay is shared for two nodes to exchange their information to increase the resource efficiency [4–7]. It is helpful to use the relay efficiently in the limited number of relays [7].

In two-way relay, the transmission time (as related to frame length) should be managed efficiently. Because the total transmission time is limited, the total transmission time must be split up for each transmission of users and relay. Therefore, unnecessary transmission time allocation brings insufficiency of transmission time for other nodes. It decreases the total transmitted data of the system [7]. In addition, the allocation of classical relaying cannot be applied directly to two way relaying because of its different operational protocol.

This paper was supported by Korea Institute of Science and Technology Information (KISTI).

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 363–368, 2020.
https://doi.org/10.1007/978-981-13-9341-9_62

In the previous research, we presented a new frame length allocation to maximize average data rate for two-way relay systems under adaptive modulation [7]. However, the general case of transmission time allocation is also required without the adaptive modulation. In addition, we only proposed the method to find the optimal value by numerical approach without the exact optimal value in the previous research. In this paper, we proposed the optimal time allocation as the closed form without the adaptive modulation assumption.

2 System Model

Similar to [7], the two-way relay system is described in Fig. 1.

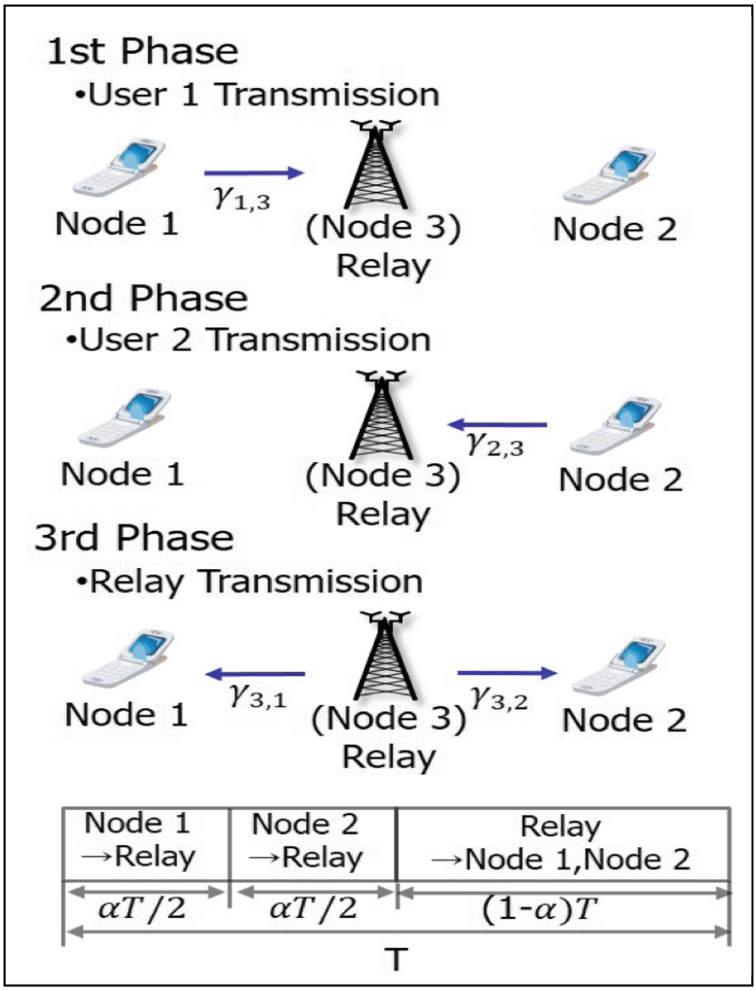


Fig. 1. System model of two relaying system

We assumed 3 node scenario of two-way relaying. Each user who want to exchange information is defined as node 1, node 2. The relay is defined as node 3. The signal to noise ratio (SNR) from node i to node j is defined as $\gamma_{i,j}$. The each fading channel is assumed to be Rayleigh fading channel.

In the first and second phase, each user (referred as node 1, 2) transmits data to the relay. Lastly, at the third phase, the relay combines and transmits the received data to each user. It means that two-way relaying can save the time resource because the classical relaying requires four phases, where user 1 transmits data to the relay, the relay transmits data to user 2, user 2 transmits data to the relay and the relay transmits data to user 1. We define the total transmission time as T. The each allocated time for phase 1 and phase 2 is defined as $\alpha T/2$. We assume that the same transmission time is allocated to the transmission for node 1 and 2 for simplicity. Therefore, the allocated time for phase 3 is defined as $(1-\alpha T)$. We define α. the time allocation factor.

3 Proposed Time Allocation

3.1 Total Average Transmission Data Rate

First we analyzed the total average transmission data rate of two-way relay system by the probability of successful transmission and the target data rates.

If we define the target rate from node 1 or 2 to the relay as R, then we can obtain the required SNR from node 1 or 2 to the relay ($\Gamma_{1,3}$ *or* $\Gamma_{2,3}$) to support R as follows [3, 8]:

$$\Gamma_{1,3} = \Gamma_{2,3} = 2^R - 1. \tag{1}$$

Similarly, because of the superposition, the target rate from the relay to node 1 and 2 as 2R, and the required SNR from the relay to node 1 and 2 as follows [3, 8]:

$$\Gamma_{3,1} = \Gamma_{3,2} = 2^{2R} - 1. \tag{2}$$

Then, we define the outage probability for each link as $\mathrm{P}_{1,3} = \Pr(\gamma_{1,3} < 2^R - 1)$, $\mathrm{P}_{2,3} = \Pr(\gamma_{2,3} < 2^R - 1)$, $\mathrm{P}_{3,1} = \Pr(\gamma_{3,1} < 2^{2R} - 1)$ and $\mathrm{P}_{3,2} = \Pr(\gamma_{3,2} < 2^{2R} - 1)$. Because we assume the Rayleigh fading channel model, the outage probability for each link as follows [3, 8]:

$$\mathrm{P}_{i,j} = \begin{cases} 1 - \exp\left(-\frac{2^R-1}{\bar{\gamma}_{i,3}}\right), i = 1,2 \\ 1 - \exp\left(-\frac{2^{2R}-1}{\bar{\gamma}_{3,j}}\right), j = 1,2 \end{cases}. \tag{3}$$

$\bar{\gamma}_{i,j}$ is the SNR from node i to node j.

The average successful transmission data rate of each link can be obtained as follows:

$$\bar{R}_{i,j} = \begin{cases} R\left(1 - \mathrm{P}_{i,3}\right), i = 1,2 \\ R\left(1 - \mathrm{P}_{3,j}\right), j = 1,2 \end{cases}. \tag{4}$$

Note that the transmission from the relay to node 1 or node 2, SNR is required to support $2R$. However, the actual transmission data rate is R because of superposition coding.

Then, the total system average transmission data is proportional to the allocated time. Because of decode and forward relaying protocol, it is determined by the weakest link [7]. The total average transmission data is as follows [7]:

$$\bar{R}_{Total}(\alpha) = \mathrm{MIN}\left(\frac{\alpha T \bar{R}_{1,3}}{2}, (1-\alpha) T \bar{R}_{3,2}\right) + \mathrm{MIN}\left(\frac{\alpha T \bar{R}_{2,3}}{2}, (1-\alpha) T \bar{R}_{3,1}\right), 0 \le \alpha \le 1. \tag{5}$$

$MIN(a, b)$ means the minimum value between a and b.

Then, we should determine the optimal time allocation to maximize the total average transmission data [7].

$$\underset{\alpha}{\mathrm{MAX}}\, \bar{R}_{Total}(\alpha), \ \textit{subject to}\, 0 \le \alpha \le 1. \tag{6}$$

3.2 Optimal Time Allocation

Since the functions are the linear functions related to α in MIN function of (5), the functions in MIN function are also concave functions by the definition of the concave function.

Similarly, the sum of minimization of concave functions is easily proven to be concave function.

It means that (5) is concave function and we can the potential points of inflection by solving the following equation:

$$\frac{\alpha T \bar{R}_{1,3}}{2} = (1-\alpha) T \bar{R}_{3,2}, \ \ \frac{\alpha T \bar{R}_{2,3}}{2} = (1-\alpha) T \bar{R}_{3,1} \tag{7}$$

Therefore, the potential points of inflection are as follows:

$$\alpha_1 = \frac{2(1-P_{3,2})}{3 - P_{1,3} - 2P_{3,2}} = \frac{2}{\exp\left(\frac{2^{2R}-1}{\bar{\gamma}_{3,2}} - \frac{2^{R}-1}{\bar{\gamma}_{1,3}}\right) + 2},$$

$$\alpha_2 = \frac{2(1-P_{3,1})}{3 - P_{2,3} - 2P_{3,1}} = \frac{2}{\exp\left(\frac{2^{2R}-1}{\bar{\gamma}_{3,1}} - \frac{2^{R}-1}{\bar{\gamma}_{2,3}}\right) + 2}. \tag{8}$$

The optimal time allocation is one of (8). We can the optimal time allocation to maximize (5) with α_1 *or* α_2.

4 Simulation Results

We verify the superiority of the proposed time allocation in Fig. 2. The total average data rate of the proposed time allocation is shown with respect to the normalized SNR, compared with the total average data rate of the optimal time allocation, which is obtained by the numerical approach. We define the normalized SNR as the SNR which is not applied by the path-loss term. The path loss exponent is assumed to be 4. The target rate R is assumed to be 1. The relay is positioned at 1/4, 1/2 and 1/3 from the node 1 to the node 2. The result shows that the proposed time allocation maximized the total average data rate as the optimal time allocation in Fig. 2.

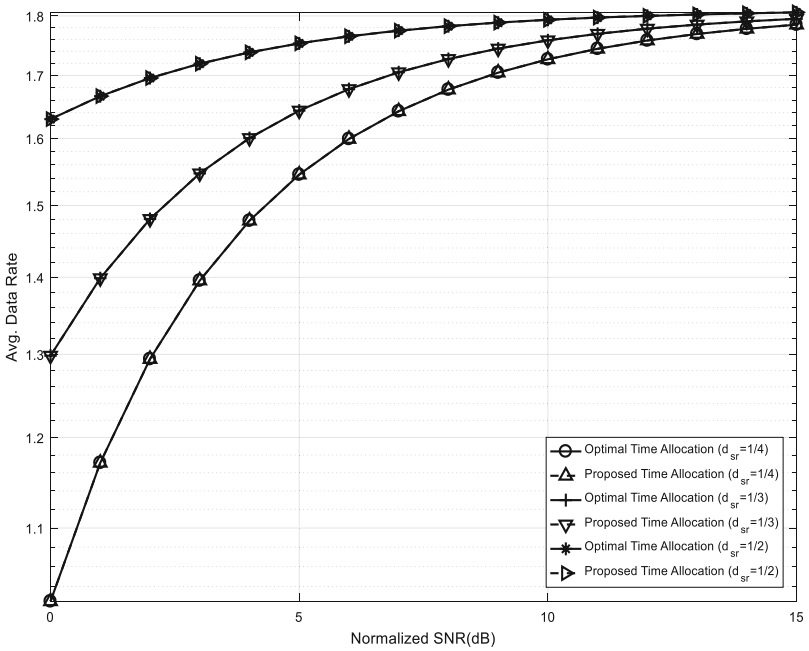


Fig. 2. Total average data rate with respect to the normalized SNR

5 Conclusion

The optimal time allocation is derived as the closed form to maximize the total average transmission rate of two-way relay system. For this purpose, the total average transmission data rate is derived by the probability of successful transmission and the target rate when the two-way relay system is used. Based on the derived total average transmission data rate, we proposed the optimal time allocation for two-way relay system. The simulation result verifies that the proposed time allocation is optimal.

References

1. Pabst, R., Walke, B.H., Schultz, D.C., Herhold, P., Yanklomeroglu, H., Mukhergee, S., Viswanathan, H., Lott, M., Zirwas, W., Dohler, M., Aghvami, H., Falconer, D.D., Fettwels, G.P.: Relay-based deployment concepts for wireless and mobile broadband cellular radio. IEEE Commun. Mag. **42**(9), 80–89 (2004)
2. Cho, J., Haas, Z.J.: On the throughput enhancement of the downstream channel in cellular radio networks through multihop relaying. IEEE J. Sel. Areas Commun. **22**(7), 1206–1219 (2004)
3. Kwon, T., Lim, S., Choi, S., Hong, D.: Optimal duplex mode for DF relay in terms of the outage probability. IEEE Trans. Veh. Technol. **59**(7), 3628–3634 (2010)
4. Jitvanichphaibool, K., Zhang, R., Liang, Y.C.: Optimal resource allocation for two-way relay assisted OFDMA. IEEE Trans. Veh. Technol. **58**, 3311–3321 (2009)
5. Chen, M., Yener, A.: Power allocation for F/TDMA multiuser two way relay networks. IEEE Trans. Wireless Commun. **9**, 546–551 (2010)
6. Han, Y., Ting, S.H., Ho, C.K., Chin, W.H.: Performance bounds for two-way amplify-and-forward relaying. IEEE Trans. Wireless Commun. **8**, 432–439 (2009)
7. Kwon, T.: A new frame length allocation to maximize average data rates for two-way relay system. In: Proceedings of International Conference on Computing Convergence and Applications (ICCCA) 2016, December 2016
8. Kwon, T., Lim, S.: Optimal duplex selection for decode and forward relay systems with power allocation. KSII Trans. Internet Inf. Syst. **10**(12), 5910–5923 (2016)

Effective Data Transfer Method Using Local Network in Building IoT Environments

Hwirim Byun[1], Hyeyoung Kang[1], Hyun-Woo Kim[2], and Young-Sik Jeong[2(✉)]

[1] Allforland Inc., Seoul, South Korea
{hazzzly, heyzero}@all4land.com
[2] Department of Multimedia Engineering, Dongguk University, Seoul, South Korea
{hwkim, ysjeong}@dongguk.edu

Abstract. A range of studies regarding the development of basic technologies for IoT in buildings are under way because the use of cloud services with this type of system is being more widely applied. If the IoT environment in a building is established using an existing cloud service, the client has to receive data from the cloud through a wide area network (WAN). Focusing on the fact that sensor data can be received directly from the hub through a local area network (LAN), this study proposes a method applying a WAN and LAN in parallel. This new approach is expected to increase the data reception speed on the client's end in the IoT environment found in buildings.

Keywords: Internet of Things · Parallel network · Monitoring system · Cloud computing

1 Introduction

The Internet of Things (IoT) is being applied in an increasingly wider range of areas, and a growing number of cases are applying sensor-network based IoT clouds for buildings. IoT used in buildings provides services according to the state of the building based on a review and analysis of sensor and CCTV data. A number of sensors and CCTV systems are used to cover a building of immense size, which inevitably entails both cloud and big data technologies. Cloud services used as basic system technologies, such as Amazon Web Service (AWS) and Microsoft Azure, are commonly used in many different facilities owing to their high efficiency in large streaming data analysis and storage. Figure 1 shows a scheme indicating the connection between an IoT in a building and its clients through an existing cloud service. If multiple administrators in the building access the data in real time under this type of setting, the access of multiple administrators may cause a delay in the data transmission because the amount of real-time data input is greater than the network speed provided by the cloud server. To resolve this problem, this study proposes Distributed Data Parallel Receive (DDPR), a method providing a smooth real-time data streaming service using the internal local network of the building to facilitate the simultaneous access of multiple administrators within the building. IoT in buildings is chiefly utilized by internal administrators for

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 369–375, 2020.
https://doi.org/10.1007/978-981-13-9341-9_63

building management, disaster response, and monitoring, among other purposes. Therefore, the parallel use of a local area network (LAN) for receiving sensor data and a wide area network (WAN) for receiving cloud analysis data will increase the performance of the system.

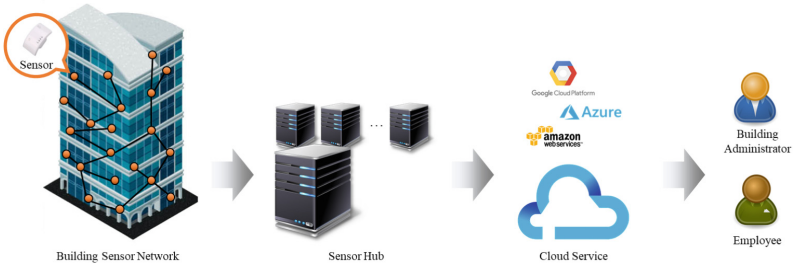


Fig. 1. Usual building IoT scheme using cloud service

2 Related Works

The method proposed in this study applies a concept similar to the existing cloud data off-loading method.

Flinn [1] and Balan [2] suggested a data distribution method in which a single central server stores the current resources on each device. The device distributes data to devices with the most available resources by referring to a database containing information such as the current availability and CPU load of the server.

Marinelli [3] and Huerta-Canepa [4] suggested a real-time data distribution method and scenario for services such as real-time media provisioning and museum applications.

Deboosere [5] presented a protocol used to search the most efficient hub node when the location of a client device changes. This method minimizes the server delay and quickens the response time, supporting the client mobility and minimizing the decline in performance.

Cuervo [6] took a step further from the data distribution designed for improved data processing efficiency, and proposed a data distribution method intended for energy saving of the device.

3 Distributed Data Parallel Receive (DDPR)

Using a WAN and a LAN in parallel depending on the situation, the presented method aims to increase the general speed of the data reception and ensure a fast response. Our method operates only in environments where a device used for data access uses the same network as the sensor hub. Thus, it is applicable to settings such as an IoT cloud environment of a single building in which users access data through the same network, for instance in building monitoring. Previous research has focused on data distribution

techniques, compression and environment configurations. This paper uses a method considering the specificity of the network environment.

Figure 2 shows the DDPR scheme. The sensors and CCTV form a sensor network based on a LAN, and send data to the nearest hubs on the network. The hubs store and manage data transmitted from CCTV and send them to the cloud through a WAN. With the existing data utilization method using a cloud, the client accesses the cloud through a WAN and receives data from the cloud, which is represented by the black arrow in Fig. 2.

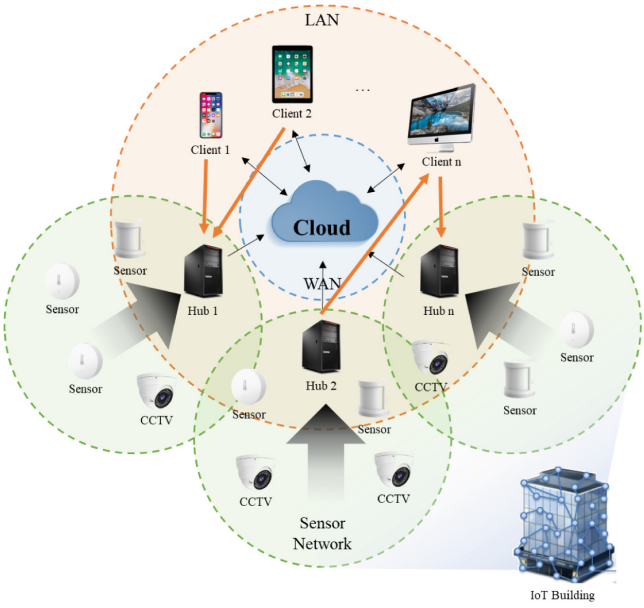


Fig. 2. Scheme of DDPR and legacy IoT environments

The use of DDPR is indicated by the orange arrow in the same figure. The client receives the required data through a LAN when requiring certain data. This approach secures a wider network bandwidth compared to a conventional method that receives data from the cloud solely through a WAN.

Figure 3 shows the order of the DDPR process, which is the process flow when the client requests data access to the cloud server.

(1) The client sends a request message to the server to request data. A request message consists of the StartTime of the requested data, the EndTime of the requested data, the DataID representing the ID of the requested data, and the ClientAddress, which shows the IP address of the client requesting the data.
(2) The cloud server sends a ForwardingMessage to the hub containing the requested data based on the DataStoreTable to request a data transmission to the client. The DataStoreTable is a table produced when the data collected from the hub are

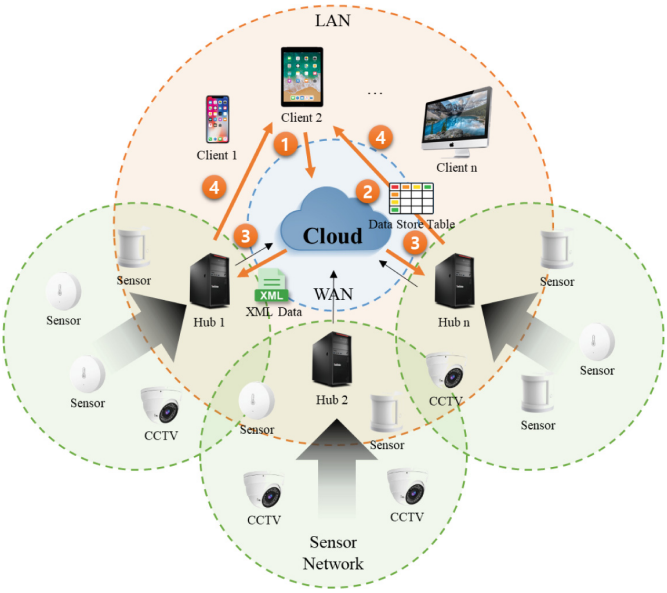


Fig. 3. Procedure of DDPR

received. It consists of the StartTime, indicating the start time of the data stored on each hub; the EndTime, indicating the end time of the data; DataID, representing the ID of the data; and the HubAddress, which is the IP address of the hub storing the data. The cloud server can identify the hub storing the data required by the client through the DataStoreTable. A ForwardingMessage is a message that commands the data transmission from the hub to the client. It consists of the StartTime, indicating the start time of the data; EndTime, indicating the end time of the data; DataID, representing the ID of the data; and ClientAddress, which is the IP address of the client.

(3) The hub that receives a ForwardingMessage delivers the requested data to the ClientAddress referring to the DataID. If there are two or more types of requested data, it can receive multiple ForwardingMessages. In addition, a single ForwardingMessage can be delivered to multiple hubs, and data can be transmitted to a single client. In this way, the client is able to receive data without interruption even if the data were recorded on multiple hubs owing to a hub change in the sensor network.

(4) The hub transmits the data requested by the client to a LAN. If the EndTime is defined in the RequestMessage sent by the client, the session is closed after the requested data are sent. When the EndTime is undefined, the data are transmitted in real time until an end message is received from the client.

4 Performance Evaluation

Figure 4 shows a comparison of the time consumed when data are directly received from the popular commonly used cloud service Amazon AWS, and the time consumed when data are received from the hub using DDPR. This study measured the latency and file transmission speed until the moment the data are received from the server and hub through the client. The file sizes used were 374 KB, 1.5 MB, and 100 MB, and Amazon AWS was configured for standard storage and non-encryption. The time was delayed in sending lightweight data because the response time increased as a result of using DDPR. However, the transmission speed was improved, and the sending time for bigger files was reduced.

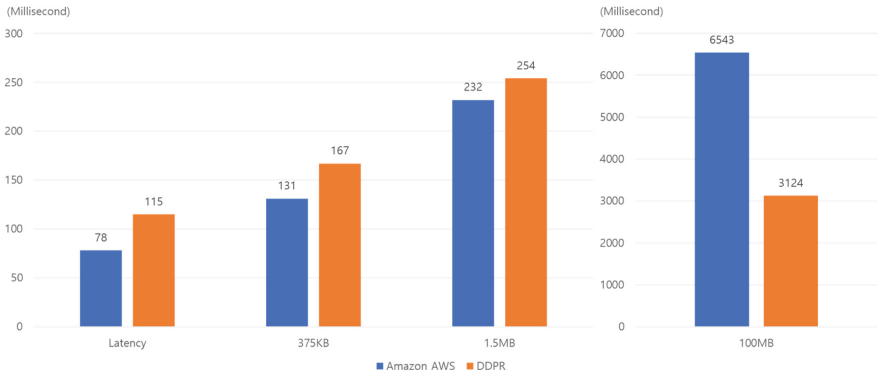


Fig. 4. Small size file download speed comparison between Amazon AWS and GDPR

Figure 5 shows the results of the measured compound file transmission speed according to the DDPR application. Here, 50% of the compound file is stored on the hub, and 50% on the cloud server. Therefore, it was defined as a parameter to measure the efficiency of the parallel data transmission through DDPR. Compound files of 100 MB, 500 MB, and 1 GB in size were used for the test. The data transmission speed improved by 35% when DDPR was used as compared to the conventional method. Because the messaging between the cloud server and hub takes time during the initial period, the improvement in speed increases proportionately with the data size.

In sending large CCTV images, or with sensor streaming, the frequency of the initial messaging is low, and the amount of transmitted data is large, thereby directly improving the performance.

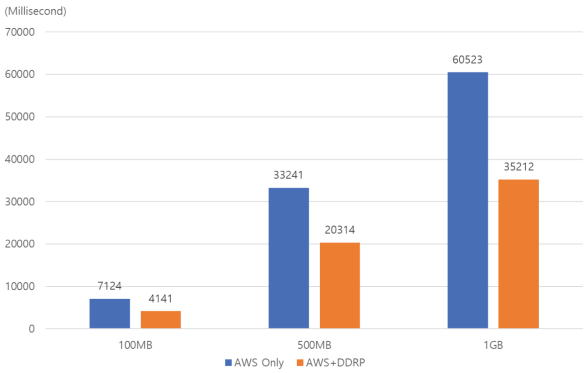


Fig. 5. Big size file download speed comparison between Amazon AWS Only and AWS + DDPR

5 Conclusion

This study proposed the use of DDPR, which achieves an improved performance by allowing a client, connected to the same network as the hub in a parallel LAN and WAN environment, to use the LAN and WAN in parallel. A parallel data transmission is realized through a standardized XML transmission, which communicates data transmittable on a LAN directly between devices without going through the cloud server. Our DDPR minimizes the network streaming that occurs when using an external cloud server, achieving a fast response time and wide bandwidth. DDPR is applicable to building monitoring environments in which sensor data uploaded onto each millisecond should be immediately collected and addressed. DDPR is limited in that it operates only in specific environments within a narrow scope. However, it will also offer some positive effects such as a reduction in cloud costs for an entire building, as well as a reduced data reception speed and response time. In future studies, an application scenario will be established to investigate the effects of the improved speed demonstrated in the present study in a real environment.

Acknowledgements. This research was supported by a grant (17NSIP-B135746-01) from the National Spatial Information Research Program (NSIP) funded by the Ministry of Land, Infrastructure and Transport of the Korean Government.

References

1. Flinn, J., Park, S., Satyanarayanan, M.: Balancing performance, energy, and quality in pervasive computing. In: The 22nd International Conference on Distributed Computing Systems, Austria, Vienna, 2–5, July 2002, pp. 217–226 (2002)
2. Balan, R., Satyanarayanan, M., Park, S., Okoshi, T.: Tactics-based remote execution for mobile computing. In: The 1st International Conference on Mobile Systems, Applications and Services, San Francisco, California, 05–08, May 2003, pp. 273–286 (2003)

3. Marinelli, E.E.: Hyrax: cloud computing on mobile devices using MapReduce. Masters thesis, Carnegie Mellon University, pp. 1–107, September 2009
4. Huerta-Canepa, G., Lee, D.: A virtual cloud computing provider for mobile devices. In: MCS 2010 Proceedings of the 1st ACM Workshop on Mobile Cloud Computing and Services: Social Networks and Beyond, San Francisco, California, 15 June 2010, pp. 1–6 (2010)
5. Deboosere, L., Simoens, P., Wachter, J.D., Vankeirsbilck, B., Turck, F.D., Dhoedt, B., Demeester, P.: Grid design for mobile thin client computing. Future Gener. Comput. Syst. **27**(6), 681–693 (2011)
6. Cuervo, E., Balasubramanian, A., Cho, D.-K., Wolman, A., Saroiu, S., Chandra, R., Bahl, P.: MAUI: making smartphones last longer with code offload. In: MobiSys 2010 Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services (2010)

Facial Photo Recognition Using Deep Learning in Archival Record Management System

Gantur Togtokh[1], Kyung Chang Kim[1(✉)], and Kang Woo Lee[2]

[1] Department of Computer Engineering, School of Engineering, Hongik University, Seoul, Korea
gantur.t@gmail.com, kckim@hongik.ac.kr

[2] Department of Computer Engineering, School of Engineering, Halla University, Wonju, Korea
kwlee@halla.ac.kr

Abstract. A lot of information is stored in the archival record management system (archive), including photos and pictures. It is important to intelligently organize photos in the archive. Current approaches use face recognition technology based on deep learning to manage photos. However, due to the rapid growth of the volume of photos in the archive, face recognition processes on large photoset take lots of processing time. In addition, low resolution photo with small faces in the archive is difficult to identify and recognize. In this paper, we propose a method to identify and retrieve facial photos from the archive. In our approach, photo metadata is used for searching photos in the archive, and resolution enhancement step based on DCSCN model is used to reconstruct photos of low resolution to high resolution. Experiment shows that the proposed approach can search and retrieve facial photo quickly from large photoset and is efficient for identifying small faces of low resolution photos.

Keywords: Record management system · Face recognition · Deep learning · Photo metadata

1 Introduction

Due to rapid growth of the volume of photo data, the need to organize photo data in the archival record management system (archive) intelligently has been an important issue in recent years. At present, photo management systems use face recognition technology based on deep learning. Typical face recognition [1] steps include face detection, face alignment, face encoding, and face prediction. Searching photos in large photoset directly by the face recognition steps consumes a lot of time.

In addition, low resolution photos with small faces in the archive cannot be correctly recognized (or identified) by typical face recognition processes. To address these two problems, we propose an approach which uses photo metadata and resolution enhancement step, and develop a large photo archive using the proposed approach. In our approach, to handle the time consuming problem on large photoset, we generate photo metadata for each photo using the results of face recognition processes in advance, and then use it to search photos in the archive. To handle the difficulty in

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 376–380, 2020.
https://doi.org/10.1007/978-981-13-9341-9_64

recognizing low resolution photo with small faces, we reconstruct low resolution photo to high resolution before applying face recognition processes.

There are number of methods to enhance the resolution of image (Image Super Resolution). Recently, deep learning based methods has been becoming more popular. Many deep learning based methods such as DCSCN [2], SRCNN [3] and DRCN [4] are proposed. DCSCN [2] model is used in the resolution enhancement step in our approach. This model achieves state-of-the-act performance and is faster and more efficient in computation [2].

For photo management, there are many researches and commercial systems related to it. In Pixelsior [5], a platform service design for photo management is proposed. It provides a unified view of photos to all application and eliminates the complexity of metadata and photo editing management. A mobile photo management system named Photo4W is developed in [6]. Key technologies in the Photo4W are face annotation and science classification. An android-based plant identification system is developed in [7]. In its feature extraction step, important features such as PHOG and color are extracted and fused to form the final feature space. In [8], the implementation of photo editing system and live wallpaper based on android platform is proposed.

Most photo management systems are implemented for mobile systems and do not covers the difficulty issue to recognize low resolution photos in its photoset. Our proposed system is implemented for desktop and cloud systems because the archive consists of a great deal more photos (i.e. order of magnitude increase in volume) than personal photo management systems.

The rest of this paper is organized as follows. Section 2 covers our proposed photo archive management system. Section 3 shows the experiment results of our proposed system on a large photoset. Section 4 concludes the paper.

2 Proposed System

Our proposed system has six main components: Register photo, archive new photos, generate photo metadata, train model, recognize faces in input photo and search photos. In this section, we introduce the main features used in our proposed system. To search photos quickly in large photoset, our proposed system generates photo metadata using the results of face recognition processes on the photo. In the our proposed system, search photos module searches photos from photo metadata database and retrieve photo from photos archive using results of the search.

2.1 Face Recognition

As we mentioned in the introduction, face recognition has four basic steps: face detection, face alignment, face encoding, and face prediction. We added one more step namely “Enhance resolution”. Some old photos in archive can have low resolution. In case of low resolution photos, small faces in such photos cannot be recognized well by typical face recognition steps.

Enhance resolution step firstly checks the resolution of the photo. If the photo is of low resolution, it reconstructs low resolution photo to high resolution one. Steps

sequence of our face recognition is shown in Fig. 1. To reconstruct low resolution photo to high resolution one, we used DCSCN [2] model which is one of the deep learning based image super resolution methods.

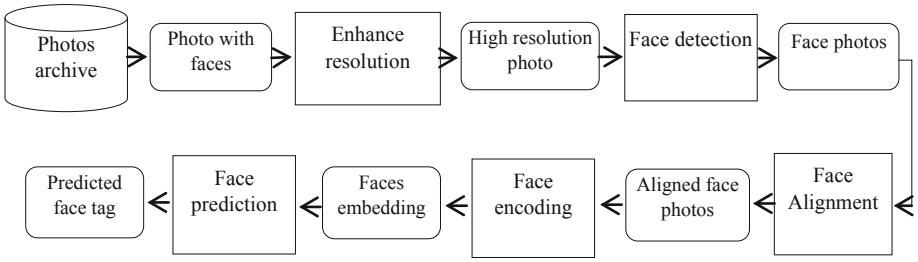


Fig. 1. Step flow diagram of face recognition processes in our proposed system.

2.2 Generating Photo Metadata

To generating the photo metadata, our system uses results of the face recognition processes. The results are unique identification of the face (face embedding), predicted photo tag (name of the person), face location (coordinates of face rectangle), etc. These photo metadata are stored in photo metadata database. Figure 2 shows the flow chart to generate the photo metadata. Photo metadata database schema in our proposed system has a relation named "face_metadata" as shown in Fig. 3. The face metadata relation's attributes are explained in Table 1.

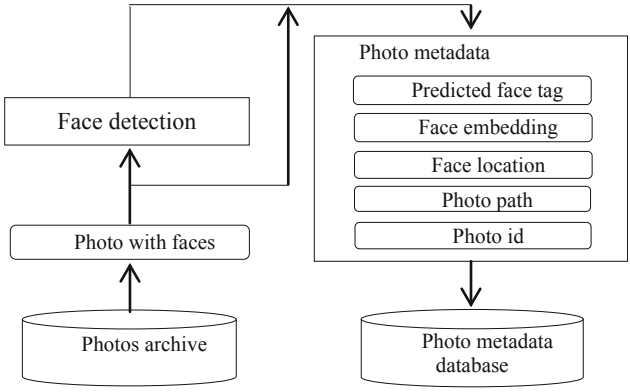


Fig. 2. Flow chart of generating photo metadata

```
face_metadata ( M_ID integer, PHOTO_ID integer, PHOTO_PATH string,
                PREDICTED_TAG string, FACE_LOCATION_TOP integer,
                FACE_LOCATION_LEFT integer, FACE_LOCATION_RIGH integer,
                FACE_LOCATION_BUTTON integer,FACE_EMBEDDING string );
```

Fig. 3. Photo metadata database schema.

Table 1. Description of attributes of face_metadata relation

Attributes	The description
M_ID	The key of value for photo metadata
photo_id	Unique photo number
photo_path	Path of photo in archive system
predicted_tag	Tag value of predicted face (name of person)
face_location_top	Y ordinate of face rectangle's pixel 1
face_location_left	X ordinate of face rectangle's pixel 1
face_location_right	X ordinate of face rectangle's pixel 2
Face_location_bottom	Y ordinate of face rectangle's pixel 2
Face_embedding	Face's unique identification feature

3 Experiment

3.1 Test Environment

Specification of the test environment is as follows. The computer used in our test is Intel® core i7-3370, CPU is 3.5 GHz, 11 GB RAM, and GTX 1070Ti. The Operating system is Ubuntu 18.04 LTS. We also used Tensorflow and dlib libraries. Our large photo archive management system was written using python programming language.

3.2 Test Results

We measured consumed times for finding and listing photos that contain same face with input tag both from photos archive and photo metadata database. Table 2 shows comparison of both the consumed times on a photoset with 1000 photos. Linear searching method (1: N compressions) without any sort or indexing is used for finding. Our experiment clearly shows that the consumed time for finding photos from photo metadata database is dramatically lower than direct searching from the photos archive. This is because the time for face recognition processes on each photo are eliminated.

We also calculated the percentages of correctly recognized faces on a photoset with 50 low resolution photo (with lower than 250 pixel width) using face recognition (with 3 times photo scale up) either with or without Enhance resolution step. Table 3 shows percentage of correctly recognized faces in photos in the photoset either with or without Enhance resolution step. It can be seen that the accuracy of recognizing faces correctly improves from 81% to 93%.

Table 2. Consumed times for finding photos from photo metadata database and photos archive

Photoset	Consumed time to find photos from photos archive	Consumed time to find photos from photo metadata database
1000 photos	93 s	2 s

Table 3. Percentage of correctly recognized face either with or without enhance resolution step

Photoset	Correctly recognized faces (without enhance resolution step)	Correctly recognized faces (with enhance resolution step)
50 photos	81%	93%

4 Conclusion

Recently, using face recognition technology based on deep learning has become more popular for photo management. In the case of large photosets, searching photos directly in photos archive by face recognition consumes lots of processing time because of face recognition processes on each photo. Test result for searching photos in achieve has showed that using photo metadata which is composed of results of face recognition processes dramatically reduces searching time. In addition, to solve difficulty in recognizing small faces in low resolution photos using typical face recognition processes, we used enhance resolution step based on the DCSCN model. Test result for recognizing low resolution photo has showed that applying the step in face recognition processes improves the accuracy of recognizing low resolution photos. We believe that the large photo archive system implemented using our approach can be used to managing any large photo sets efficiently.

Acknowledgments. This work (Grants No. S2601476) was supported by project for Cooperative R&D between Industry, Academy, and Research Institute funded by Korea Ministry of SMEs and Startups in 2018.

References

1. Modern Face Recognition with Deep Learning. https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78
2. Yamanak, J., Kuwashima, S., Kurita, T.: Fast and accurate image super resolution by deep CNN with skip connection and network in network. In: 24th International Conference of Neural Information Processing, pp. 217–225 (2017)
3. Dong, C., Loy, C.C., He, K., Tang, X.: Learning a deep convolutional network for image super-resolution. In: European Conference on Computer Vision, pp. 184–199 (2014)
4. Kim, J., Lee, J.K., Lee, K.M.: Deeply-recursive convolutional network for image superresolution. In: Computer Vision and Pattern Recognition, pp. 1637–1645 (2016)
5. Kyungho, J., Sharath, C., Karthik, D., Ko, S.Y.: Pixelsior: photo management as a platform service for mobile apps. In: Proceedings of 8th USENIX Workshop on Hot Topics in Storage and File Systems (HotStorage 16), pp. 1–5 (2016)
6. Sun, F.M., Li, H.J., Wang, X.M.: Photo 4W: mobile photo management on what, where, who and when. Neurocomputing **119**, 59–64 (2013)
7. Zhao, Z.Q., Ma, L.H., Cheng, Y.M., et al.: ApLeaf: an efficient android-based plant leaf +identification system. Neurocomputing **151**, 1112–1119 (2015)
8. Choi, K., Ihm, S.Y., Park, Y.-H.: An implementation of a photo editing system and live wallpapers on Android platform. In: Proceedings of 7th International Conference on Embedded and Multimedia Computing (EMC) Embedded and Multimedia Computing Technology and Service, pp. 715–720 (2012)

A Dynamic Plane Scaling Method for Smart Dust Environments

Joonsuu Park and KeeHyun Park(✉)

Computer Engineering Department, Keimyung University,
Daegu 42601, Republic of Korea
parkjoonsuu@gmail.com, khp@kmu.ac.kr

Abstract. In a smart dust environment, throughout which small devices are spread to communicate with each other, the unique properties of the smart dust make it difficult to use a typical dual plane network structure. Therefore, we propose in this paper a dynamic scaling method that allows the resources of control and data plane to fluctuate organically. Our proposed method achieves this goal by unifying various variables into critical values, introducing the concept of a neutral (available) plane, and using predictive algorithms.

Keywords: Internet of Things · IoT · Smart dust · Dual plane · Dynamic scaling

1 Introduction

An IoT (Internet of Things), in which various smart devices (sensors and actuators) communicate and converge, is together with artificial intelligence, machine learning and big data one of the most significant IT technologies of current era [1]. We are already using IoTs for logistics and medical services today, and are actively trying to utilize the IoT concept in a range of other fields, such as 'smart' city and 'smart' home applications [2, 3]. In addition, IoT can be extended beyond the static fixed device layout to extreme environments, or areas where people cannot actually or artificially determine the location of the device [4].

1.1 Smart Dust

Smart dust technology can detect and manage the information of surrounding temperature, humidity, acceleration and pressure through a wireless network by spraying dust sensors that are very small in size in a physical space such as building, a road, or on clothes [5, 6]. In [5], the key elements of the MEMS technology of Smart Dust were reviewed and the research challenges were explained. In [6], a MEMS technology was used to pack an autonomous sensing, computing, and communication system into a cubic-millimeter mote. The mote can be the basis of integrated, massively distributed sensor networks. Smart dust has also attracted the attention of the military, as it can be sprayed from an airplane and used to collect a variety of information about areas that are not accessible to humans [7].

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 381–384, 2020.
https://doi.org/10.1007/978-981-13-9341-9_65

In a typical network environment in which human intervention is involved, sensors can be installed in optimal positions at the required positions, so that various patterns of data generated or processed within the network can be predicted, controlled and managed. However, the placement characteristics of scattered smart dust create various idiosyncrasies for different reasons; for example, the sensors may be clustered in a specific area, or may have difficulty with communication [5]. The following are examples of some of the various idiosyncrasies specific to a smart dust network.

- Uneven amount of measurement data
- Uneven amount of measurement data processing demand
- Random measurement data generation time
- The frequency of occurrence of different measurement data

In a static IoT environment, the positions of the gateways can be uniformly arranged, or placed at optimized positions as needed, and the ratio of the control plane and the data plane can be maintained at an optimal state. However, in a smart dust environment with various data pattern idiosyncrasies, a static approach can reduce efficiency, or in the worst-case scenario make the system inoperable.

We are planning to apply the Smart Dust concept to IoT systems. As the first step of our plan, a dynamic scaling method that allows the resources of control and data plane to fluctuate organically is proposed in this paper.

1.2 DPDK (Data Plane Development Kit)

Developed by Intel, the DPDK (Data Plane Development Kit) is a high-speed packet processing library that can run on various CPU architectures [8]. The DPDK is a collection of data plane libraries and NIC (network interface controller) drivers for high-speed packet processing. The DPDK provides a programming framework for Intel x86 processors and enables the rapid development of high-speed data packet networking applications.

While a typical network driver is tailored to the Ethernet interface required by the operating system, the DPDK provides an interface to directly control the hardware without interfering with the operating system. It also includes a memory management library and a data structure library that can handle high-speed packet processing [9].

The DPDK supports development tools that can process data at high speeds. But in a smart dust environment where data patterns are irregular, static plane scales cannot be determined.

1.3 Dual Plane in SDN

The separation of the control and data planes in the SDN (Software Define Network) inevitably simplifies the function of the switch. That is, the control plane becomes simpler and the data plane becomes more complex [10]. This means that most of the resources of the gateway are used by the data plane. However, it differs in special circumstances such as smart dust. The control plane, which determines the connection and path, plays a larger role in the smart dust environment, where sessions frequently connect and disconnect. However, the features of the data plane are not reduced.

When a large amount of data is generated and a large amount of measurement data is poured into the gateway, the data plane requires more resources than the control plane. Therefore, the resource ratio between the control plane and the data plane, which is fixed in a static environment, must be organically scaled down/up in a smart dust environment.

2 Dynamic Plane Scaling

In general, the simplest way to think about dynamically scaling the control/data plane is by specifying a threshold. If a session connection that exceeds the threshold is requested, the number of control planes is increased. Conversely, when a session connection below a threshold is requested, the number of control planes is reduced. The same is true for the data plane. However, there is a problem with how much resources should be allocated.

2.1 Available Planes

To solve the problem of allocating system resources, a proper unit of resource is needed. As shown in Fig. 1, the available plane area has a resource group named unit plane.

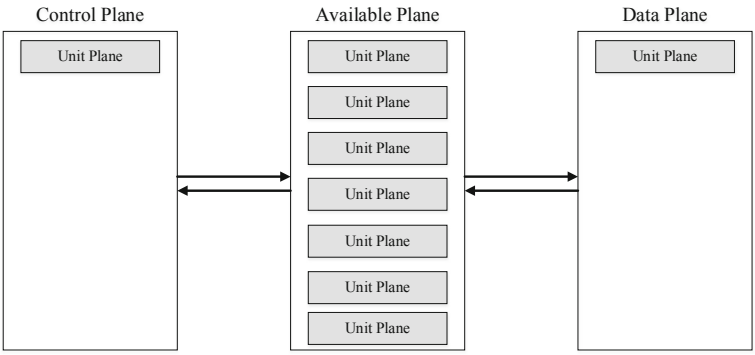


Fig. 1. Dynamic plane scaling based on the available plane

This resource group (unit plane) is assigned to the place where it is requested on a group basis.

2.2 Threshold Prediction

The threshold value can be set statically, but the dynamic nature of smart dust makes it difficult to choose correctly. In addition, when the dust device leaves the network for physical or environmental reasons, a threshold value should be set to reflect that. Because a static threshold is difficult to consider, a prediction algorithm for the values is required.

A smart dust environment cannot reproduce the same environment. Therefore, the algorithm to use for prediction cannot be labeled. This means that we need to consider reinforcement learning.

3 Conclusion

We have proposed a dynamic plane scaling method that can organically scale the network plane in the very unusual and unique network environment specific to smart dust.

Dynamic plane scheduling of dual planes reduces many variables to one variable (threshold) and predicts them. It also reduces the resources for plane expansion/shrinking by uniting the resource units into units.

Acknowledgments. This research was supported by the Basic Science Research Programs through the National Research Foundation of Korea (NRF), funded by the Ministry of Education, Science and Technology (No. NRF-2018R1D1A1B07043982).

References

1. Sethi, P., Sarangi, S.R.: Internet of things: architectures, protocols, and applications. J. Electr. Comput. Eng. **2017**, 1–25 (2017)
2. Mao, X., Li, K., Zhang, Z., Liang, J.: Design and implementation of a new smart home control system based on internet of things. In: 2017 International Smart Cities Conference (ISC2), pp. 1–5. IEEE (2017)
3. Rahmani, A.M., et al.: Exploiting smart e-health gateways at the edge of healthcare internet-of-things: a fog computing approach. Future Gener. Comput. Syst. **78**, 641–658 (2018)
4. Farris, I., Taleb, T., Khettab, Y., Song, J.S.: A survey on emerging SDN and NFV security mechanisms for IoT systems. IEEE Commun. Surv. Tutorials **21**, 812–837 (2018)
5. Kahn, J.M., Katz, R.H., Pister, K.S.J.: Next century challenges: mobile networking for "smart dust". In: Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking, Seattle, Washington, USA (1999)
6. Warneke, B., Last, M., Liebowitz, B., Pister, K.S.: Smart dust: communicating with a cubic-millimeter computer. Computer **34**(1), 44–51 (2001)
7. Rosenthal, M.M.: Gamebits: digital tricks. Games 6 (2000)
8. Intel®: About – DPDK, 30 September. https://www.dpdk.org/about/
9. Intel®: Data Plane Development Kit 18.08.0 documentation. https://doc.dpdk.org/guides
10. Casado, M., Koponen, T., Shenker, S., Tootoonchian, A.: Fabric: a retrospective on evolving SDN. In: Proceedings of the First Workshop on Hot Topics in Software Defined Networks, pp. 85–90. ACM (2012)

Analysis of Psychological Stability and EEG-Based Control Efficacy of Infants by Stimulation Technique-Infant Car Seat Seating Environment

Jeong-Hoon Shin[1(✉)] and Hyeon-Cheol Seo[2]

[1] School of Information Technology, Daegu Catholic University, Hayang-Ro 13-13, Hayang-Eup, Gyeongsan-si, Gyeongbuk, Republic of Korea
only4you@cu.ac.kr
[2] Department of Computer Engineering, Daegu Catholic University, Hayang-Ro 13-13, Hayang-Eup, Gyeongsan-si, Gyeongbuk, Republic of Korea
tjgus7784@cu.ac.kr

Abstract. Infant car seats provide safety and comfort to infants seated in cars. Nonetheless, these seats will fail to provide comfort unless infants adapt themselves to the seats. To resolve such problems, various types of infant car seats have been rolled out in the market, which combine the effects of vibration and sound, etc. However, it is unclear whether the infants are seated in these car seats longer than in existing car seats, due to their psychological stability, anxiety, or curiosity towards such novelties. In this study, we analyzed EEG-based psychological stability through various techniques of stimulation applied by infant car seats and proposed new stimulation techniques that will allow infants to be seated in infant car seats for prolonged periods of time, in a psychologically stable state. These techniques can be used in various fields, such as neurofeedback therapy, in the period ahead.

Keywords: Infant car seat · Electroencephalogram (EEG) · Brain-activated state · Psychological stability

1 Introduction

As the ownership of private cars increases, contemporary society has witnessed a surge in fatalities arising from traffic accidents. Traffic accidents, particularly car crashes, are a major cause of unintended death and disability among infants aged 3 years or older. To prevent that, infant car seats have been legislated in most countries [1, 2].

Infant car seats, if used properly, can reduce infant injuries by about 70%, deaths by 90%, serious injury by 75%, and minor injury 67% [3–6]. However, most infants often feel uncomfortable being restrained by infant car seats, or often feel isolated from their parents, and as a result, cannot remain seated for long periods. The emotional state of these infants contributes to a reduction in the rate of car seat use in infants. To resolve these problems, various types of infant car seats with vibration and sound effects, etc., have been introduced [7]. These car seats arouse the interest of infants and thereby

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 385–392, 2020.
https://doi.org/10.1007/978-981-13-9341-9_66

increase the duration that the infants are likely to remain seated in car seats. However, it is unclear whether the infants are seated in car seats longer due to their psychological stability, anxiety, or curiosity towards such new effects.

To analyze the exact psychological state of infants, electroencephalogram (EEG) characterization analyses, which reflect human cerebral activation state, should be carried out. An EEG is an electrical signal measured non-invasively through electrodes attached to the surface of the head, to measure the electrical activity of the brain.

Among them, the bands associated with the psychological state of infants are theta and alpha waves. Theta waves, which occur mainly in the process leading to emotional stability or sleep, are more distributed in children than in adults, and have been found to be associated with various conditions such as memory, psychic powers, creativity, concentration, anxiety, etc. Alpha waves, which occur mainly in relaxed states, show an increase in their amplitude in more stable and relaxed states, and appear usually in the form of regular waves.

2 Experiment Design

2.1 Subjects

Subjects, totaling 19 infants, consisted of 10 boys and 9 girls between the ages of 4 to 11 months. Subject groups were divided into a reference group and control group for the verification of the efficacy of car seats. The reference group was comprised of 7 infants (4 boys and 3 girls) while the control group consisted of 12 infants. (6 boys and 6 girls).

2.2 Methodology of Experiment

For the reference group, electroencephalography (EEG) data were collected from the infants sitting in car seats for 15 min in a non-stimulated manner. At this time, subjects were expected to feel uncomfortable, and experience changes in their feelings over time as they were separated from their parents, and had to sit alone in the infant car seats for a long time.

These emotional changes could be compared with those observed in the control group, and were used as reference emotions that might change with time.

In the control group, EEG data were collected from the subjects sitting in the infant car seats for the first 3 min without any stimulation. Then, music stimulus (music that infants normally enjoyed) was applied for 3 min, while their EEG data were collected. For the next 3 min, the infants were again exposed to music while their EEG data were collected. For another 3 min, the infants were exposed to the stimuli of both music and vibrations varying with the music simultaneously while their EEG data were collected. For comparison with infant car seats available in the market, the vibration stimulus of ordinary patterns was applied while EEG data were collected. For the last 3 min, the EEG data of subjects were collected without applying any stimulus, in order to analyze the psychological state of the infants according to variations in time and stimuli applied. Based on that, we performed the analyses to determine the effects of music and

vibration on the psychological stability in the infants, and compared and analyzed the changes of the emotional state of the infants over time with those of the emotional state in the reference groups.

2.3 Data Collection

The Fp1 and Fp2 regions are responsible for body movements based on thoughts, plans, and judgments. In this experiment, we used a band-type EEG measurement device to minimize any possible discomfort that could be felt by infants in order to ensure that the data could be collected while they were in a state of utmost psychological stability. The data were collected from the Fp1 and Fp2 regions according to the time and method defined in Fig. 1.

In this study, we performed a two-dimensional analysis using Valence-Arousal to analyze the state of psychological stability based on the brain science approach [10]. Valence and Arousal values were mapped to each quadrant based on the results of Z-score calculation, and the emotion states represented by each quadrant were as follows Fig. 2:

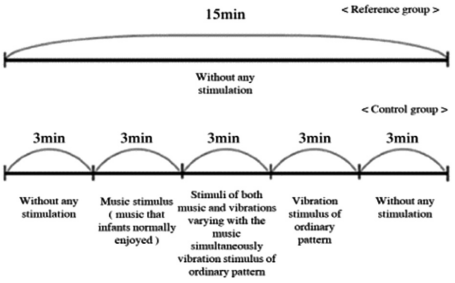


Fig. 1. Environment for collection of EEG data from infants.

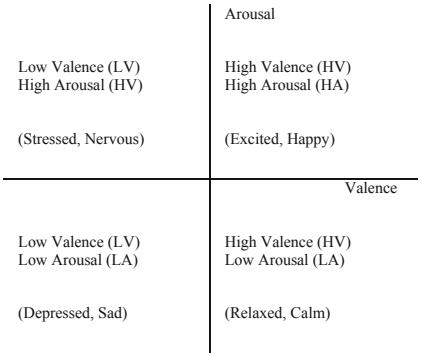


Fig. 2. Two-dimensional model of Valence – Arousal (Z score).

2.4 Data Analyses

In this experiment, we analyzed the emotional state of infants by using Valence value and Arousal value, and based on the results, compared the efficacy of conventional infant car seats and the proposed infant car seat. Valence represented an emotional dimension corresponding to the enjoyment showing the extent of pleasantness-unpleasantness as the first dimension of emotion, and was related to the alpha wave asymmetry of the prefrontal lobe in connection with the EEG characteristics [8]. If alpha waves in the left frontal lobe showed higher values than those in the right frontal lobe, the right brain was relatively activated. If the value of the alpha waves were higher in the right frontal lobe than in the left frontal lobe, the left brain was relatively activated. The activation of the right frontal lobe is associated with negative emotions,

while the activation of the left frontal lobe is associated with positive emotions. The Valence used in this study can be defined by the following formula.

$$\text{Valence} = \text{Fp1}(\alpha) - \text{Fp2}(\alpha) \tag{1}$$

Arousal corresponds to the degree of arousal that shows the extent of arousal-relaxation in the emotional experience, which is different from the Valence emotional state, such as excitement-relaxation, activation-inactivation, etc., and is classified into the most relaxed state (sleepy, relaxed) to the most aroused state (excited, angry) [9]. Arousal is associated with beta waves related to emotional states such as arousal and excitement, and alpha waves related to the emotional states such as inactivation and relaxation of the brain as manifested by the EEG signals. Therefore, it would be reasonable to consider the Arousal by taking the ratio of alpha waves and beta waves together into consideration if the level of arousal is to be assessed. Thus, in this study, we intended to define the Arousal based on the ratio of alpha to beta waves according to the formula below:

$$\text{Arousal} = \frac{\alpha\text{waves}}{\beta\text{waves}} = \frac{\text{Fp1}(\beta) + \text{Fp2}(\beta)}{\text{Fp1}(\alpha) + \text{Fp2}(\alpha)} \tag{2}$$

3 Analyses of the Experiment Results

3.1 Quantitative EEG Analysis

In this study, we performed fast Fourier transforms (FFT) on the EEG signals collected from the infants, and carried out quantitative EEG analyses. To observe and analyze the ratio of theta and alpha waves among the frequency components of the EEG signals generated at this time, and determine the psychological state of the infants, we performed statistical analyses based on the changes in relative energy of alpha and theta waves and Z-score over time.

3.1.1 Changes in Relative Energy of Alpha Waves over Time

The results of analysis, shown in Fig. 3, show the average energy of the alpha waves from 7 infants the reference group and 12 infants in the control group measured every 3 min.

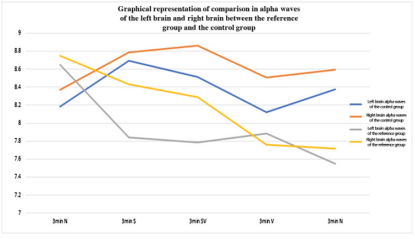


Fig. 3. Changes in the relative energy of alpha waves over time

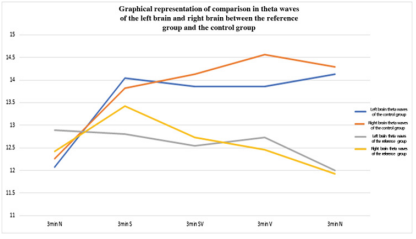


Fig. 4. Changes in the relative energy of theta waves over time

The cerebral activation state of the reference group was analyzed by using the Valence value mentioned in Sect. 2.4. The results showed that the alpha wave value of the left brain was higher than that of the right brain, although the psychological state was initially positive. This indicates that the right frontal lobe is activated over time and that the psychological state changes to a negative or depressed state.

In the control group, by contrast, the alpha wave value was higher in the right brain than in the left brain, suggesting that the left frontal lobe was activated. Moreover, the control group was found to maintain a positive psychological state over time, considering that the activation of the left frontal lobe was associated with positive emotions.

3.1.2 Changes in Relative Energy of Theta Waves over Time

The results of analysis, shown in Fig. 4, show the average energy of the theta waves from 7 infants in the reference group and 12 infants in the control group measured every 3 min.

In order for the psychological state to be changed to a relaxed and stable state, this was considered to have occurred when the energy values of alpha and theta waves were relatively dominant.

As shown in Figs. 3 and 4, the energy values of alpha and theta waves were found to gradually decrease with time in the reference group. This is in contrast to the results of the control group. The reference group was found to be in an anxious and emotionally stressed state, rather than a psychologically relaxed and stable state over time.

In the control group, the energy values of alpha and theta waves were found to increase gradually over time. This suggests that the control group came to be in a relaxed and psychologically stable state over time.

3.2 Analysis of Psychological Stability Index

When Valence and Arousal are represented as two-dimensional models, Quadrant 1 represents an excited and happy emotional state (HV+HA). Quadrant 2 represents a stressed and nervous emotional state (LV+HA). Quadrant 3 represents a sad and depressed emotional state (LV+LA), and Quadrant 4 represents a relaxed and calm emotional state (HV+LA). In this study, the emotional state of infants sitting in the infant car seats was analyzed by using the Valence-Arousal two-dimensional model to compare the efficacy of the newly-proposed infant car seat.

3.2.1 Analysis of Psychological Stability Index in the Reference Group

Figure 5 shows the results of the Z-score analysis on one of the infants in the reference group over time. As can be seen in Fig. 5, the Z-score of the reference groups trends was distributed more in Quadrant 2, showing an angry and stressed emotional state and Quadrant 3 showing a sadder and more depressed emotional state than in Quadrant 1, which showed an excited and happy emotional state, and Quadrant 4 which indicated a relaxed and calm emotional state. This suggests that the infants who sat in the infant car seats became angry, stressed, and depressed over time.

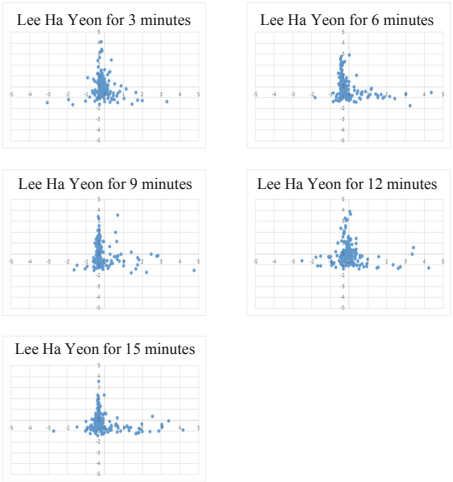


Fig. 5. Results of Z-score analysis on one of them infants in the reference group over time

3 minutes		6 minutes		9 minutes	
47.7	31.9	41.0	38.7	42.0	36.3
54.4	46.0	48.7	51.6	52.3	49.4

12 minutes		15 minutes	
34.9	44.0	37.4	40.9
45.4	55.7	52.4	49.3

Fig. 6. Mean Z-score of infants in reference group over time (7 infants)

Figure 6 shows the distribution of mean Z scores of each quadrant based on the Valence-Arousal of 7 infants in the reference group over time. As can be seen in Fig. 6, infants in the reference group exhibited a high frequency of Z-score distribution in Quadrant 2 and Quadrant 3, representing a negative emotional state, which suggests that they did not feel stable psychologically while sitting in the infant car seats.

3.2.2 Analysis of Psychological Stability Index in the Control Group

Figure 7 shows that the frequency of Z-score distribution was higher in Quadrant 1 and Quadrant 4 due to the influence of music and vibration stimulation over time in the control group. This shows that the infants came to be in a happy and psychologically stable state over time, in response to music and vibration stimuli.

Figure 8 shows the distribution of mean Z-score of each quadrant for 12 infants of the control group, based on the Valence-Arousal over time. As shown in Fig. 8, the mean Z-score of infants in the control group was higher in Quadrant 1 and Quadrant 4, representing the positive emotional state over time, whereas Quadrant 2 and Quadrant 3 represented a negative emotional state.

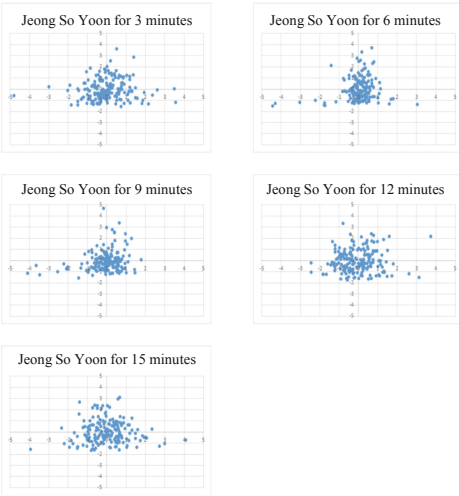


Fig. 7. Results of Z-score analysis on one of them infants in the control group over time

3 minutes		6 minutes		9 minutes	
30.2	47.2	34.8	45.9	34.8	43.3
50.1	52.5	45.5	53.8	47.8	54.1

12 minutes		15 minutes	
30.6	45.7	26.5	51.6
47.9	55.8	47.1	54.8

Fig. 8. Mean Z-score of infants in control group over time (12 infants)

4 Conclusion and Direction for Succeeding Studies

The quantitative EEG analysis of infants in the reference group showed that the alpha wave values of the left brain were higher than those of the right brain over time, leading to activation of the right frontal lobe. In addition, the energy values of alpha and theta waves, the indicators of anxiety and emotional stress, decreased with time. The Z-score value also showed an increased distribution in Quadrant 2 and Quadrant 3, representing their anxiety and depression over time. Therefore, it can be found that the psychological state of subjects in the reference group was changed to an anxious and emotionally-stressed state over time, after they were seated in the infant car seats.

In the control group, the emotional state of infants sitting in infant car seats changed to a happy and stable state over time. That was demonstrated by quantitative EEG analysis and Z score values. The quantitative EEG analysis of the control group showed that the alpha wave values were higher in the right brain than in the left brain over time, leading to activation of left frontal lobe. Additionally, the energy values of alpha and theta waves, the indicators of a stable and relaxed psychological state, were found to have increased. This shows that the subjects in the control group exhibited a psychological state of happiness and stability gradually. Z-score values were also found to be distributed more in Quadrant 1 and Quadrant 4, representing happiness and emotional stability over time. Therefore, subjects in the control group were found to feel happiness and emotional stability over time in response to the musical stimuli preferred by these subjects, concurrent vibration stimulus varying with their preferred musical stimuli, and vibration stimuli of certain patterns after they were seated in the infant car seats.

In particular, the musical stimuli preferred by these subjects and concurrent vibration stimulus varied with their preferred musical stimuli (stimulus of music

preferred by infants plus stimulus of vibration varying with the music) led the indicators of happiness and emotional stability to increase by more than 10.4% and 13.7%, respectively, in the control group, compared to the reference group in the same time slot.

Acknowledgments. This work was supported by the sabbatical research grant from Daegu Catholic University in 2015.

This experiment was conducted in compliance with the research ethics rules, with the consent of the subject's legal representative, with sufficient explanation regarding the experiment.

References

1. Kroeker, A.M., Teddy, A.J., Macy, M.L.: Car seat inspection among children older than 3 years: using data to drive practice in child passenger safety. J. Trauma Acute Care Surg. **79**(3 Suppl 1), S48–S54 (2015)
2. Lee, J.-H., Lee, Y.-H.: Safety seats research in toddler and preschooler parents. Korean Parent-Child Health J. **10**(2), 110–112 (2007)
3. Vaca, F., Anderson, C.L., Agran, P., Winn, D., Cheng, G.: Child safety seat knowledge among parents utilizing emergency service in a level 1 trauma center in Southern California. Am. Acad. Pediatr. **110**(5), 61–65 (2002)
4. Ramsey, A., Simpson, E., Rivara, F.P.: Booster seat use and reasons for non-use. Am. Acad. Pediatr. **106**, E20 (2000)
5. Louis, B., Lewis, M.: Increasing car seat use for toddlers from inner-city families. Am. J. Public Health (AJPH) **87**, 1044–1045 (2011)
6. Park, H.-G.: Measures for improvement and diffusion of domestic car seats for infants and children. Korean Soc. Des. Cult. **14**(1), 135–148 (2008)
7. Takada, J.: Child Safety Seats for Vehicles. U.S. 04979777, 25 December 1990
8. Choppin, A.: EEG-Based Human Interface for Disabled Individuals: Emotion Expression with Neural Networks. Master Thesis, Information Processing, Tokyo Institute of Technology, Yokohama, Japan (2000)
9. Ramirez, R., Palencia-Lefler, M., Giraldo, S., Vamvakousis, Z.: Musical neurofeedback for treating depression in elderly people. Front. Neurosci. **9**, 1–10 (2015)
10. Ha, J.-M., Park, S.-B.: Assessment of color effect on the indoor color schemes and illuminance change – focused on prefrontal EEG alpha and beta signal analysis. J. Archit. Inst. Korea **10**, 57–65 (2017)

Parallel Graph Clustering Based on Minhash

Byoungwook Kim[1(✉)], Jaehwa Chung[2], Joon-Min Gil[3], and JinGon Shon[2]

[1] Department of Computer Engineering, Dongguk University, Gyeongju, Korea
bwkim@dongguk.ac.kr
[2] Department of Computer Science, Korea National Open University, Seoul, Korea
{jaehwachung, jgshon}@knou.ac.kr
[3] School of Information Technology Engineering, Catholic University of Daegu, Daegu, Korea
jmgil@cu.ac.kr

Abstract. Graph clustering is a technique for grouping vertices having similar characteristics into the same cluster. It is widely used to analyze graph data and identify its characteristics. Recently, a large-capacity large-scale graph data is being generated in a variety of applications such as a social network service, a world wide web, and a telephone network. Therefore, the importance of clustering technique for efficiently processing large capacity graph data is increasing. In this paper, we propose a clustering algorithm that efficiently generates clusters of large capacity graph data. Our proposed method efficiently estimates the similarity between clusters in the graph using Min-Hash and generates clusters according to the calculated similarity. In the experiment using real world data, we show the efficiency of the proposed method compared with the proposed method and existing graph clustering methods.

Keywords: Graph clustering · Spark

1 Introduction

The graph consists of vertices and edges and has been regarded as an important data structure by modeling the relationship between the vertices constituting the graph. Graph data is used in various fields such as social network (SNS), telephone network, bio-network, world wide web (WWW) and road network [1]. Among the various techniques for analyzing graph data and extracting meaningful information, graph clustering is a technique that measures the similarity of the vertices constituting the graph and classifies the entire graph into multiple clusters so that the vertices related to each other belong to the same cluster, And graph clustering is an important technique used in various applications such as social network analysis and community detection [2], image segmentation, and protein-protein correlation.

Recently, a large amount of graph data with large number of vertices and trunks has been generated due to low price of storage media, activation of social networks, development of web technology, and high availability of various data [1]. Large-scale graph data consists of tens of small to large billions of vertices and corresponding

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 393–395, 2020.
https://doi.org/10.1007/978-981-13-9341-9_67

trunks. For example, the World Wide Web (WWW) contains 50 billion web pages and more than one trillion links, and the social network Facebook also contains more than 800 million peaks and over 100 billion friends. As the size of the generated graph data increases, the clustering methods proposed in the past have a problem in that the time required for clustering a large amount of graph data increases (Fig. 1).

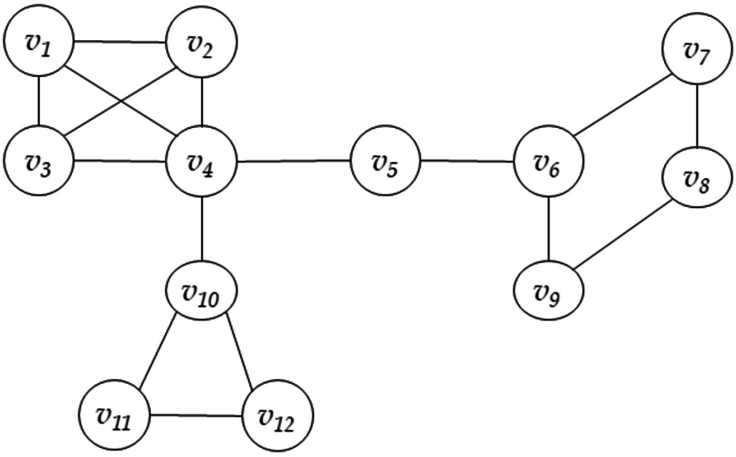


Fig. 1. A simple graph

In this paper, we use Min-Hash, a kind of locality sensitive hashing (LSH) technique, to approximate similarities between vertices in a non-directional and non-weighted large- If the similarity between vertices is greater than the user-defined threshold, the two vertices are merged into the same cluster. We also propose a technique to calculate the similarity between clusters in which vertices are merged and to merge two clusters if the similarity between clusters is larger than the user - defined threshold.

2 Related Works

2.1 Minhash

The Minhash technique is a kind of Locality Sensitive Hash (LSH), which approximates how two sets are similar. The similarity of two sets can be defined as Jaccard similarity [9] as in Eq. (1).

$$\mathrm{Sim}(\mathrm{S},\mathrm{T}) = (\mathrm{S} \cap \mathrm{T})/(\mathrm{S} \cup \mathrm{T}) \tag{1}$$

The Min-Hash technique approximates Jaccard similarity of two sets S and T, and the basic principle is to map the set S to one of the elements belonging to the set S. To map a set to an element, a hash function is used in the Min-Hash technique. The value

mapped to one element is called the Min-Hash value, and all the elements belonging to the set S are substituted into the hash function, and the smallest value is the Min-Hash value. At this time, the probability that the Minhash value of both sets S and T is the same is equivalent to Jaccard similarity.

3 Parallel Graph Clustering

In this paper, we propose a method to efficiently calculate similarity using Min-Hash and to cluster graph data based on it. The first step is the preprocessing step (line 3 in Fig. 3) to generate signatures to perform clustering. The second step (line 4 in Fig. 3) performs clustering using the signatures generated in the preprocessing step.

Various clustering techniques have been proposed to extract meaningful information from graph data. However, the conventional techniques require many operations in order to determine the degree of similarity between vertices, and thus clustering can not be efficiently performed in a large amount of graph data.

Initially, each cluster is considered to contain only one vertex, and it is judged whether or not it can be merged into the same cluster through calculation of similarity between clusters and clusters.

4 Conclusions

In this paper, we propose an efficient graph clustering technique using Min-Hash. Since the min-hash technique approximates Jaccard similarity and the similarity between vertices in graph data can be represented by Jaccard similarity, we propose a technique to efficiently calculate similarities between vertices using Min-Hash. In addition, we show that the merged clusters can effectively generate the signatures of the merged clusters. Experiments show that the proposed algorithm performs clustering at a high speed, while the quality of clusters shows good results.

References

1. Kang, U., Faloutsos, C.: Big graph mining: algorithms and discoveries. ACM SIGKDD Explor. Newslett. **14**(2), 29–36 (2012)
2. Newman, M.E.J., Girvan, M.: Finding and evaluating community structure in networks. Phys. Rev. E **69**(2), 026113 (2004)

Golf Swing Recognition Using Low-Cost Smart Insole

Eun-Young Lee, Jinu Kim, and Dongho Kim(✉)

Department of ICMC Convergence Technology, Soongsil University, 369, Sangdo-ro, Dongjak-gu, Seoul, Republic of Korea
{ella, kjwl008}@gsclab.kr, cg@su.ac.kr

Abstract. The purpose of this study was to develop a golf swing recognition system using forty pressure sensors embedded insoles and a smartphone. This system detects golf swing event using pressure data from a pair of insoles and recorded sound level. And mobile application provides the center of pressure trajectory overlaid at swing events. Users can record the pressure data and observe center of pressure trajectory of their golf swing easily with this system. In the future, this system could be a new golf swing analysis system in the field.

Keywords: Golf swing · Swing analysis · Smart insole · Plantar pressure

1 Introduction

Golf is one of the most popular recreational sports and there are millions of golfers in the world. Most golfers want to use technologies that analyze golfer performance to develop their skills. The golf swing is a complex movement, which is influenced by the golfer's feet movement [1], so common golf swing analysis methods include 2D video image [2, 3], 3D motion capture [4, 5] and pressure distribution analysis [6, 7]. However, these technologies are expensive and complex for amateur golfers. Due to this situation, research and development of wearable device-based golf swing analysis technology which can support both professional golfer and amateur golfer are being carried out.

Systems using pressure sensor to recognize and analyze the swing of a golfer are divided into a shoe-type wearable device system and an insole-type wearable device system. Conventionally developed shoe-type devices limit golf shoes which is the golfer's main equipment, and insole-type devices also limit the natural golf swing due to thickness or flexibility problems. In addition, the number of pressure sensors of both shoe-type and insole-type are less than five so that cannot make more detailed recognition and analysis.

The objective of this paper is to describe the design of insole-type wearable device system including dozens of pressure sensor which can be used to assist the player in developing a correct swing. And we also demonstrate golf swing recognition with our newly developed system.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 396–400, 2020.
https://doi.org/10.1007/978-981-13-9341-9_68

2 Golf Swing Recognition System

2.1 System Overview

The system comprises a pair of pressure sensor embedded insole and a smartphone. First, when a golfer starts golf swing, each insole starts to send the data from pressure sensors via Bluetooth to the smartphone. And then, during the golf swing, the smartphone detects the event of golf swing based on collected pressure and audio data. Lastly, when the swing is finished, separate the phases according to the detected golf swing event.

Figure 1 illustrates the configuration of the system and sensor array of each insole.

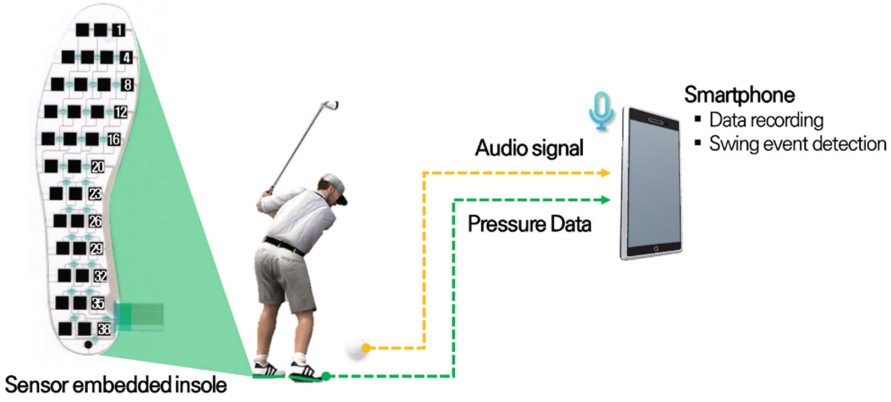


Fig. 1. Configuration of the system and insole sensor array

The sensor embedded insole is a pair of wireless insoles which consists of a mini electronic module and a pressure sensor array. A pressure sensor array consists of a flexible and thin membrane keypad which arrange forty-sensors in the shape of the sole. The sensor array is used to obtain the high-resolution pressure data from feet during the golf swing. A mini electronic module transmits pressure data converted from voltage signals into resistance respectively to the smartphone.

The mobile application allows the golfer to record the data from the sensor embedded insole and to observe the trajectory of center of pressure (COP).

2.2 Golf Swing Phase

For golf swing recognition, we divided golf swing into the following three phases:

Backswing. Backswing was from the address to the top of the backswing.

Downswing. Downswing was from the top of the backswing to the impact.

Follow-Through. Follow-through was from the impact to the finish.

2.3 Golf Swing Event Detection

The COP of each foot in the medial-lateral (COP_x) and anterior-posterior (COP_y) directions based on pressure data from forty-sensors of each insole according to the following formulas:

$$COP_i = \left[COP_x, COP_y\right] = \left[\frac{\sum_{j=1}^{40} x_j p_j}{\sum_{j=1}^{40} p_j}, \frac{\sum_{j=1}^{40} y_j p_j}{\sum_{j=1}^{40} p_j}\right] \quad (1)$$

where p_j is measured pressure and x_j and y_j are coordinate of each sensor.

The total COP is defined as the sum of right and left COP vectors and using the following equations:

$$COP_{total} = \frac{p_r}{p_r + p_l} COP_r + \frac{p_r}{p_r + p_l} COP_l \quad (2)$$

where p_r and p_l are sum of the pressure and COP_r and COP_l are COP coordinates from right insole and left insole.

To detect each golf swing event, the detection conditions were defined as follows:

Address. The address is when the swing begins.

Backswing Top. The backswing top is the point at which COP_y decreases and then begins to increase again after the address.

Impact. The impact is when club heads hit a golf ball and generate a large amount of audio signals.

Finish. The finish is the end of the golf swing.

Figure 2 shows the golf swing events and phases used in this swing recognition system.

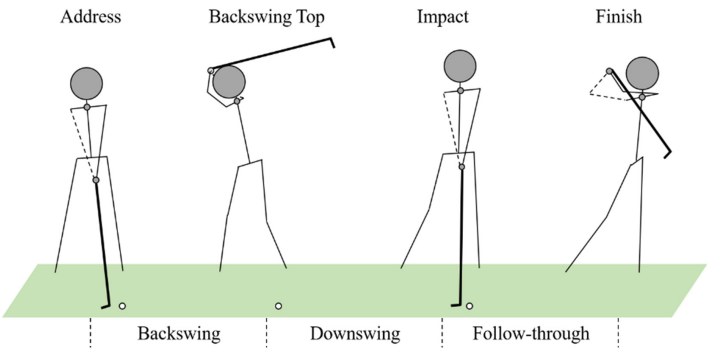


Fig. 2. The event and phases of the golf swing

3 Result

We developed a mobile application for android platform. The GUI-based application allows the golfer to record the data from the sensor embedded insole and to observe the trajectory of COP without difficulty.

When press "Start" button of the application, each insole starts to send the data from pressure sensors via Bluetooth to the smartphone and the application starts collecting the coordinates of COP from the insole until the end of golf swing. All data are obtained in real-time at the rate of 20 frames per second. And then, during the golf swing, the smartphone detects the event of golf swing based on received pressure data and sound level. Lastly, when the measurement is finished, the application separates the phases according to the detected golf swing event and displays the COP trajectory overlaid at swing events.

An amateur golfer wearing the sensor embedded insole performed golf swings for demonstration and we detected the events with the developed application.

Figure 3 shows a result screen of the application.

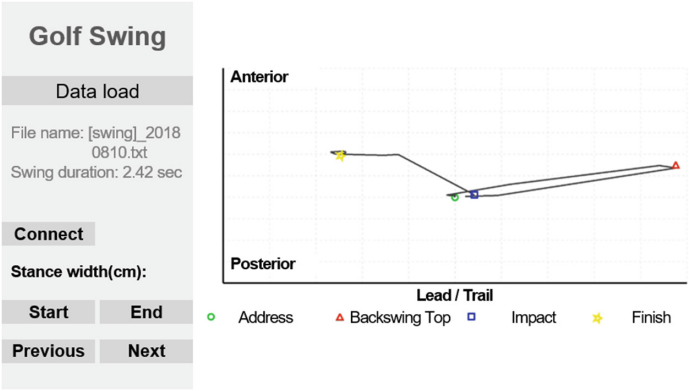


Fig. 3. Screenshot of the mobile application

4 Conclusions and Future Work

The goal of this study was to develop and demonstrate a wearable system for golf swing recognition. We used pressure data from membrane keypads and audio signal from a microphone of smartphone to detect golf swing events.

In previous studies, the insole-type golf swing analysis systems had less than five sensors due to problems such as costs, while this system which using membrane keypads has forty sensors in each insole. Additionally, a GUI-based mobile application enhances user convenience in data collecting and swing event detection. In our study, we acquired data at the 20 frames per second. To allow for a detailed and robust analysis, noise filtering and algorithm improvement are needed as well as increasing data collection frequency. To be used as an analytical tool in the field, the analysis of

the weight bearing ratio of the lead foot and the trail foot, the analysis of weight bearing ratio of the forefoot and rearfoot and the pressure distribution visualization at each phase of the golf swing are required. In addition, combining this system with the camera function of the smartphone can provide golfers both weight-bearing analysis and video analysis. We will conduct experiments with professional golfers and amateur golfers to assess validity, reliability and usability of this system.

In the future, this system could provide inexpensive and convenient swing analysis environment for both professional and amateur golfers.

Acknowledgement. This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2018-2018-0-01419) supervised by the IITP (Institute for Information & communications Technology Promotion).

References

1. Woolley, T.W., Johnson, H., Fleisig, G.S., Barrentine, S.W: Ground reaction forces and torques of professional and amateur golfers. In: Science and Golf II, pp. 58–67. Taylor & Francis (2002)
2. Urtasun, R., Fleet, D.J., Fua, P.: Monocular 3D tracking of the golf swing. In: Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 932–938. IEEE (2005)
3. Karliga, I., Hwang, J.N.: Analyzing human body 3-D motion of golf swing from single-camera video sequences. In: Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 5, pp. 493–496. IEEE (2006)
4. Egret, C.I., Vincent, O., Weber, J., Dujardin, F.H., Chollet, D.: Analysis of 3D kinematics concerning three different clubs in golf swing. Int. J. Sports Med. **24**(6), 465–470 (2003)
5. Nesbit, S.M.: A three dimensional kinematic and kinetic study of the golf swing. J. Sports Sci. Med. **4**(4), 499–519 (2005)
6. Hurrion, P.: A biomechanical investigation into weight distribution and kinematic parameters during the putting stroke. Int. J. Sports Sci. Coaching **4**(1), 89–105 (2009)
7. Ball, K., Best, R.: Centre of pressure patterns in the golf swing: individual-based analysis. Sports Biomech. **11**(2), 175–189 (2012)

Blockchain Based Integrated Authentication System

Jeong Hoon Jo and Jong Hyuk Park(✉)

Department of Computer Science and Engineering, Seoul National University of Science and Technology (SeoulTech), Seoul 01811, Korea
{jojeong3766, jhparkl}@seoultech.ac.kr

Abstract. Due to the development of hardware and the development and optimization of network, the 4th industrial revolution has grown up with information and communication technology, and new technological innovations such as artificial intelligence, internet of things, quantum computing and blockchain are emerging as core driving force. The Fourth Industrial Revolution is developing towards the connection, sharing and decentralization of nodes. It is the blockchain that includes all three characteristics. The blockchain is used in many industrial fields based on the characteristics However, there is a lack of research on the integrated authentication management blockchain. In this paper, we propose an integrated authentication management blockchain using blockchain. The existing authentication system is constituted independently, and the user must suffer inconvenience of generating redundant data in a system requiring authentication. In this paper, personal information is integrated and stored in a blockchain, personal information is protected by zero-knowledge proofs, and fast response speed can be provided.

Keywords: Authentication · Blockchain · Zero-knowledge proof

1 Introduction

Recently, the development of the blockchain has led to the use of industries in various fields. Gartner has designated the blockchain as a strategic technology in 2018 [1], and predicted that by 2025, the business value of the blockchain will exceed $176 billion and will surpass $3 trillion by 2030 [2]. Blockchains are typically used in financial, medical, contents, logistics, etc. to enhance efficiency by converging with various industrial sectors. Such blockchain applications have the advantage of maintaining the latitudinal and integrity of the data and the technology to ensure the reliability of the data. This paper proposes a user's integrated Authentication system. Existing authentication systems are operated independently due to their unique characteristics, and independent authentication systems create inconvenience for users to have duplicate authentication. To address this problem, a system that inputs user data into the blockchain and manages it collectively is proposed. Data stored in the blockchain cannot be converted into data written in transactions even if accessed by malicious attackers, and data stored in the blockchain is encrypted and archived. Encrypted data is then compared/analyzed to

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 401–404, 2020.
https://doi.org/10.1007/978-981-13-9341-9_69

return true or false data when the user requests authentication based on knowledge. This allows users to authenticate anywhere and at any time.

2 Related Work

2.1 Blockchain Technology

Currency transactions were conducted by a third-party broker to increase mutual trust in the exchange of money between individuals. However, users wanted a mechanism that would allow transactions involving sensitive information to enter into transactions without third-party intervention. In 2008, Satoshi Nakamoto negated the process of being traded by a third-party intermediary and proposed a new process [3]. Bitcoin is a cryptocurrency that allows transactions to be made without interference from third parties such as financial period. It consists of a distributed P2P network that can validate/verify data from other nodes in the transaction and protect its assets. The blockchain stores user-generated data in transactions and incorporates the generated transactions to create a single block. The generated blockchain is propagated across the network to all nodes, and once the block is validated, it is connected to the existing block. Since blockchain is propagated to all other nodes, data cannot be modified. If you attempt to modify it, you must configure more than 51% of the nodes to modify the data [4]. The types of blockchains are shown in Table 1 below [5].

Table 1. Types of blockchains

Type	Public	Consortium	Private
Read Write data permission	Blockchain participants	Consortium member Blockchain participants	Specific operating entity Blockchain participants
Block creation	Blockchain participants	Consortium member	Operator
Latency	High	Low	Low
Examples	Bitcoin, Ethereum	R3 corda	Self-development

2.2 Authentication

Authentication is a system that verifies that it is true for interactions with others who are unclear. Authentication is used in a variety of areas and is typically used to increase the reliability of commodity transactions, cash transactions and security. Systems that authenticate people can usually have variety of authentication methods such as token authentication, smart card authentication, voice recognition, fingerprint recognition, and iris recognition. About 56% of the above-mentioned authentication methods are used as fingerprint recognition methods [6]. Whereas conventional authentication simply compares the original to the authentication data, recent authentication methods allow users to authenticate with digital signatures containing public keys and digital signatures from trusted, authenticated agencies.

2.3 Zero-Knowledge Proof

The Zero-knowledge proof covers the process of proving true sentences and verifies the truth without sharing information with the verifier verifying the true sentences. The node that owns the sentence to be proved is called the certifier, and the node that communicates information with the certifier in the demonstration process is called the verifier. Zero-knowledge must meet three requirements [7].

- Complete: Zero-knowledge is always true if the input is true.
- Soundness: if the input is false, the knowledge cannot be true.
- Zero-knowledge: No one else can get input.

Zero-knowledge proof solves the problem by verifying and repeating the sentence you are doing. By judging only true/false, the round is increased by the number of squares. Thus, after 10 rounds, the probability of proving a false statement is 1/1024.

3 System Environment

Existing authentication systems are operated independently due to their unique characteristics, and independent authentication systems create inconvenience for users to have duplicate authentication.

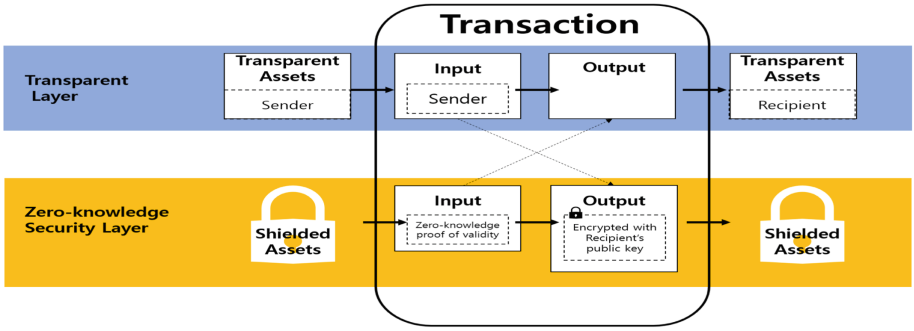


Fig. 1. Use of zero-knowledge of transaction

To address this problem, a system that inputs user data into the blockchain and manages it in an integrated manner is proposed. The proposed system model encrypts blockchain transactions, which do not expose user information, and allows verify/Check of user information stored in the blockchain. Transactions use specific Zero Knowledge proof called zk-SNARKs [8] (Fig. 1).

The following is how users store their own authentication data:

- Input of user biomechanical and authentication data
- Address generation using the user's public and private keys
- Encrypted and stored in the Transaction Data field

- Store generated addresses in 'From' field
- Electronic signature using public key to ensure user data integrity
- Gather the generated transaction sets to validate after block generation
- Create Blockchain

The way a validator requires user authentication data is shown below.

- Requesting data from a validator
- Use zk-SNARKs to prove true/false data for Authentication
- True/false return of authentication data

4 Conclusion

Recent advances in hardware and network development and optimization have begun to attract attention. Blockchains were initially used for virtual currencies but are now being used in convergence with the Internet of Things and other industries. In this study, we proposed an integrated authentication management blockchain using a blockchain. Existing authentication systems are independent, and users have experienced inconvenience in generating redundant data on independent authentication systems. In this paper, personal information was managed by integrated storage of personal information in the blockchain, and personal information was protected by permanent knowledge, and the speed of fast response was provided. Future research will implement a system that prevents the storage of malicious data when entering personal information into a blockchain.

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No 2016R1A2B4011069).

References

1. Cardellini, V., et al.: A game-theoretic approach to computation offloading in mobile cloud computing. Math. Programm. **157**(2), 421–449 (2016)
2. https://newzoo.com/insights/rankings/top-100-countries-by-game-revenues/. Accessed 29 Oct 2018
3. Nakamoto, S.: Bitcoin: a peer-to-peer electronic cash system (2008)
4. Eyal, I., Sirer, E.G.: Majority is not enough: bitcoin mining is vulnerable. In: Springer, International Conference on Financial Cryptography and Data Security, pp. 436–454 (2014)
5. Lin, I.-C., Liao, T.-C.: A Survey of Blockchain Security Issues and Challenges. IJ Netw. Secur. **19**(5), 653–659 (2017)
6. https://www.bayometric.com/global-biometric-market-analysis/. Accessed 29 Oct 2018
7. Hopwood, D., et al.: Zcash protocol specification. Technical report, 2016–1.10. Zerocoin Electric Coin Company (2016)
8. https://z.cash/blog/anatomy-of-zcash. Accessed 29 Oct 2018

Machine Learning-Based Intrusion Detection System for Smart City

Jung Hyun Ryu and Jong Hyuk Park(✉)

Department of Computer Engineering, Seoul National University of Science and Technology (SeoulTech), Seoul 01811, Korea
{jh.ryu, jhparkl}@seoultech.ac.kr

Abstract. Nowadays, the cloud computing-based Internet of Things (IoT) environment suffers from problems such as rapidly increasing data traffic volume, heterogeneity and latency. One of the typical methods to solve these problems is to utilize Fog or Edge Computing, which distributes storage and computing power concentrated in a cloud computing environment through a distributed model. However, in order to compensate for the disadvantages of this distributed network, Mist Computing has emerged as the network model closest to Internet of Things. But, there are thousands of zero-day attacks in the Internet environment of things that communicate by various protocols. Most of these attacks are small variants of previously known attacks. To effectively prevent such attacks, intrusion detection systems in the environment should be more intelligent. In this paper, in order to solve these problems, we propose an artificial intelligence-based intrusion detection system to effectively protect new or continuously changing attacks to IoT in a mist computing environment.

Keywords: Artificial intelligence · Intrusion detection · IoT · Mist computing

1 Introduction

The number of devices used in the Internet of Things (IoT) environment is expected to reach about 50 billion by 2020 [1]. New technologies are needed to stably manage and use smart devices such as mobile devices (mobile phones, tablet PCs, etc.), wireless sensors, and actuators, which are growing exponentially. In most IoT devices, that is, all devices corresponding to a network edge in the IoT environment, it is necessary to perform low latency and a large amount of data in a distributed processing method. As a result, the concept has expanded from existing Cloud Computing to Fog or Edge Computing. However, Mist Computing has emerged as a computing model that provides the closest, faster approach to those devices to effectively manage and use the explosive growth of smart devices and IoT devices. This distributed network model is located at the final stage of the network and is located nearest to the smart devices including sensors and actuators, and thus has advantages such as distributed communication, integration of heterogeneous devices, and minimization of delay time. So, Mist Computing model can solve the problems of cloud computing. However, since the mist computing model exists at the shortest distance from smart devices, the security problem is the first thing to be considered. One of the main features of the Internet

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 405–409, 2020.
https://doi.org/10.1007/978-981-13-9341-9_70

environment is that there are various protocols due to heterogeneity among the models, which is the background of thousands of Zero Day attacks. But most of these attacks are known as variants of previous types of attacks. Increasing the number of attacked objects, variants of existing attacks, that is, unknown cyber-attacks, can pose a serious crisis to the Internet environment of large scale objects. Intrusion detection in the IoT environment is different from traditional centralized cloud methods due to its special requirements - low latency, limited resources, scalability, and flexibility. Therefore, an intrusion detection system that extends to the mist computing layer should be concerned.

2 Related Work

Increasingly scaling things The internet environment has revealed the limitations of cloud computing. Hundreds of millions of Internet appliances perform communication and data processing through cloud computing, resulting in problems due to excessive traffic, high latency, and heterogeneity among models. To solve this problem, fog computing, a layered model that operates between cloud computing and endpoints, i.e., Internet devices, has emerged. This model is distributed, facilitating the distribution of services from the cloud to the Internet appliance, and consists of fog nodes. Fog Computing minimizes request and response times with supported applications and provides network connectivity and computing power for local computing resources and the required centralized cloud services for end devices. However, in order to effectively manage the continuously expanding object Internet environment, a model that operates on the object Internet layer, that is, an endpoint, has emerged. This is the mist computing model. This model is located closest to the network edge in the Internet environment to improve the efficiency of computing complex applications. When transmitting data, control is dispersed to an end node that reduces the latency of the network, thereby increasing throughput. It uses a microcomputer and a microcontroller to provide information to the fog nodes and cloud computing, and has a light computing power that exists at a close distance from the Internet appliance at the network edge.

While the cloud and fog computing models have information about the user's needs and the network global environment, the mist computing model has only information about the physical environment and the local environment, so the main purpose is to run the Internet applications of things. To this end, information about the network global environment must be transmitted to the edge, that is, the mist node.

Artificial intelligence is effectively applied to intrusion detection for cyber-attacks in the Internet environment of objects. Artificial intelligence is classified into machine learning, artificial neural network, and deep learning. Among them, Deep Learning is deepening learning stage compared to existing artificial intelligence approach, so it is less influenced by small changes and has excellent accuracy. Therefore, the deep learning is most suitable for the attack type learning of the intrusion detection system proposed in this paper.

Attack detection in intrusion detection systems through artificial intelligence is divided into signature based or anomaly based. Signature-based detection detects incoming network traffic compared to known types of attacks in the database. On the

other hand, the abnormal operation based detection method detects the attack by calculating the motion deviation from the normal traffic.

Signature-based detection methods are widely used in intrusion detection systems because they are easy to implement, have high detection accuracy, and have low false positives. However, detection rates for modified or new attacks are significantly low. On the other hand, anomaly-based detection schemes can detect a modified attack or a new attack with a high probability although the accuracy is relatively low.

It uses datasets as a means to simulate and evaluate performance of intrusion detection systems based on artificial intelligence. It is a kind of traffic collection that classifies normal and abnormal traffic that is transmitted and received in the network. The data sets considered for the intrusion detection system in this paper are the KDD'99 dataset and the NSL-KDD dataset.

The KDD'99 data set is a data set based on data collected through the existing intrusion detection system. It includes various attributes such as duration, type of protocol, and is classified into four classes [3].

The NSL-KDD dataset was created to supplement the problems of KDD'99 in 2009 and is an extended dataset of KDD'99. The KDD'99 dataset was modified to remove duplicate records and reduce the overall size of the dataset. As a result, there is a small difference in the KDD'99 dataset and attack type distribution [4].

The attack types in two datasets are classified into Normal, DoS, Probe, R2L, and U2R. All records in the dataset are categorized into the above five types and the intrusion detection system is the basis for classifying the attack type (Fig. 1).

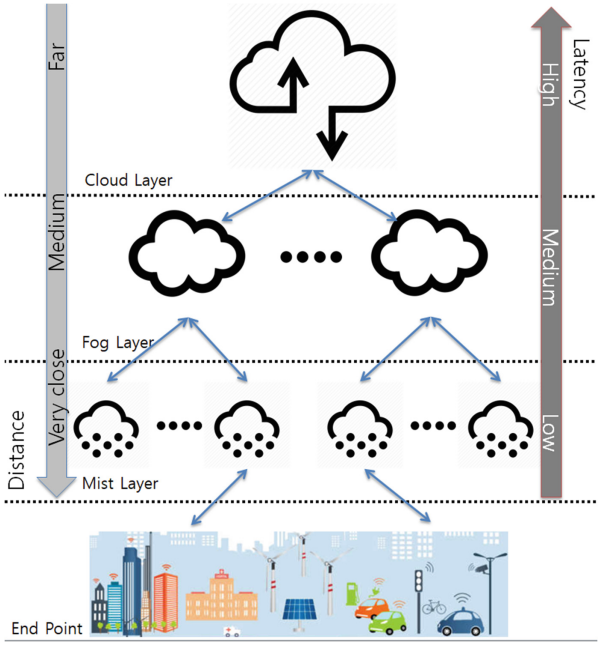


Fig. 1. Overview of mist computing

3 Proposed Method

Mist nodes are more efficient at learning and responding to attack types because they are closer to cloud and fog nodes than to smart devices in the Internet environment. Each mist node communicates with a deep learning based intrusion detection system to analyze and classify traffic. Each mist node communicates with a small object Internet smart device, so it has only local information, but analyzes the attack type by sharing learning data with neighboring mist nodes. Detection and response are faster through local learning and learning data sharing. The approach for the intrusion detection system of this paper is shown below Fig. 2.

The bottom-most endpoint layer is the layer of smart devices that actually work in the Internet environment. The middle mist layer communicates with smart devices at the nearest distance to the endpoint. The middle layer analyzes traffic from smart devices through a deep-run intrusion detection system and classifies them as normal or abnormal traffic. Each mist node continuously updates the intrusion detection system by sharing traffic received from the communicating smart devices with neighboring mist nodes. The top layer is the medium fog layer, which communicates with the cloud nodes by communicating with a specific scale of mist nodes.

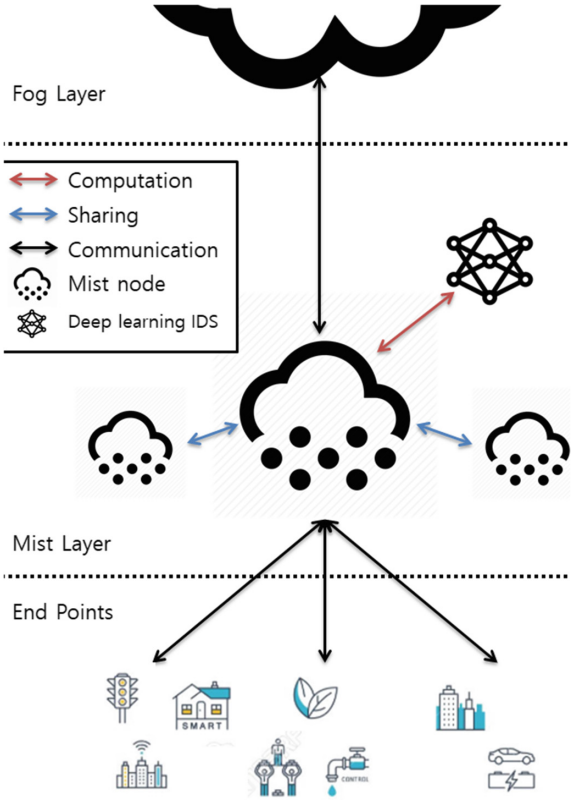


Fig. 2. Deeplearning based distributed intrusion detection system

4 Conclusion

The ever-increasing size of the Internet environment is a challenge for cloud computing. Excessive traffic, high latency, and security issues can be addressed by distributed computing approaches such as fog and mist computing. Among them, mist computing is attracting attention as the latest approach to overcome the limitations of the cloud.

Therefore, in this paper, we propose an intrusion detection system in a mist computing environment based on deep learning which is considered to be most adaptive and accurate among the artificial intelligence for the object internet environment. This can improve detection accuracy by applying a distributed intrusion detection system and enables learning and detection faster by performing learning in the nearest mist computing layer with smart devices. In the near future, we will study in detail the learning algorithms and simulations applied to the intrusion detection system presented in this paper.

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No 2016R1A2B4011069).

References

1. Iorga, M., et al.: Fog Computing Conceptual Model, pp. 500–325. NIST Special Publication (2018)
2. Diro, A.A., Chilamkurti, N.: Distributed attack detection scheme using deep learning approach for Internet of Things. Fut. Gener. Comput. Syst. **82**, 761–768 (2018)
3. Yogi, M.K., et al.: Mist computing: principles, trends and future direction. SSRG Int. J. Comput. Sci. Eng. (SSRG-IJCSE) **4**(7), 19–21 (2017)
4. Abeshu, A., Chilamkurti, N.: Deep learning: the frontier for distributed attack detection in fog-to-things computing. IEEE Commun. Mag. **56**(2), 169–175 (2018)

Computation Efficiency Analysis of Multiple GPUs and Multiple CPUs Based Cluster Computing Environments

Bongjae Kim[1], Boseon Hong[2], and Jeong-Dong Kim[1(✉)]

[1] Division of Computer Science and Engineering, Sun Moon University, Asan-si, South Korea
bjkim0422@gmail.com, kjdvhu@gmail.com

[2] Department of Computer and Electronics Convergence Engineering, Sun Moon University, 70, Sunmoon-ro 221 Beon-gil, Tangjeong-myeon, Asan-si, Chungcheongnam-do 31460, South Korea
goodcools34@gmail.com

Abstract. Multiple CPUs and multiple GPUs based cluster computing environments are popularly used and very attractive approach in various computing areas because those computing environments provide very high computing performance when compared to a typical single node based computing environments. In this paper, we compare and evaluate the performance of multiple CPUs and multiple GPUs based on cluster computing environment with MPI (Message Passing Interfaces). In the performance evaluations, we evaluate and analyze the performance of sparse matrix-vector multiply (SpMV). SpMV is one of the most widely used operations in many scientific, computational, and mathematical applications. Based on the performance evaluation results, generally, the execution time of SpMV is decreased as the number of CPUs and GPUs increase. However, there were cases that were not so in the case of GPU.

Keywords: Multiple CPUs · Multiple GPUs · High performance computing · Message passing interface · Cluster computing

1 Introduction

Recently, cluster computing environments are widely used in many computing fields such as deep learning based services and applications [1, 2]. Because multiple CPUs and multiple GPUs based cluster computing environments support relatively very high computing power when compared to a typical single node based computing system [3]. Generally, GPU provides very high performance when compared to a typical CPU because a GPU device provides more computational cores than a typical CPU. For example, NVIDIA GeForce Titan Black GPU device supports 2880 CUDA cores. However, there are memory copy overheads between main memory and global memory of GPU when GPUs are used for computations [4]. Because the necessary data must exist in the global memory of the GPU in advance. Generally, memory copy overhead increases as the number of GPUs used increase. Therefore, a policy is

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 410–415, 2020.
https://doi.org/10.1007/978-981-13-9341-9_71

required to efficiently allocate computational operations by comparing and analyzing the computational efficiency of GPU and CPU.

In this paper, we compare and analyze the performance of multiple CPUs and multiple GPUs by using sparse matrix-vector multiply (SpMV) operation. SpMV is one of the most widely used operations in many computational and mathematical applications. We can accelerate and improve the performance SpMV operation by using multiple GPUs and multiple CPUs based cluster computing environments.

The rest of this paper is organized as follows. In Sect. 2, we will show the performance evaluation results of SpMV with multiple CPUs and multiple GPUs based cluster computing environment. Finally, we conclude this paper with future works in Sect. 3.

2 Performance Evaluations

In this section, we will describe the cluster computing environments used in the performance evaluations and the results of comparing and analyzing the computational efficiency of multiple CPUs and multiple GPUs by using SpMV operation.

2.1 Evaluation Environments

Table 1 shows the performance evaluation environments. Each computing node of our cluster computing is equipped with two Intel Xeon E5-2650 CPUs and NVIDIA GeForce Titan Black. One Intel Xeon E5-2650 CPU has 8 cores. Each CPU core runs at 2.6 GHz. As we described in Sect. 1, one NVIDIA GeForce Titan Black has 2880 CUDA cores. Each GPU core runs at 889 MHz. In short, the operating speed of each core of CPU is higher than that of the GPU. The number of cores of GPU is much higher than that of CPU.

Table 1. Cluster computing environments used in the performance evaluations.

Features	Descriptions
CPU	2 × Intel Xeon CPU E5-2650 2.6 GHz v2
Memory	64 GB
OS	CentOS 7.2
Kernel Version	3.10.0
CUDA Version	CUDA 7.5
OFED Version	MLNX_OFED_LINUX 3.2
MVAPICH2 Version	MVAPICH2-2.2b
GPU	GeForce Titan Black
InfiniBand Host Channel Adapter	Mellanox ConnectX-3 VPI Adapter Dual-Port QSFP, FDR IB (56 Gb/s)

The CUDA 7.5 was installed onto our cluster computing system to manage the multiple GPUs of each computing node. MVAPICH2 2.2b [5] was used to support the communication of each compute node of our cluster computing system.

Table 2. Sparse Matrixes used in the performance evaluations.

Matrix name	Matrix size	The number of non-zero elements
road_usa	23,947,347 × 23,947,347	28,854,312
audikw	943,696 × 943,696	39,297,771
thermal	1,228,045 × 1,228,045	4,904,179
msdoor	415,863 × 415,863	10,328,339
twotone	120,750 × 120,750	1,224,224

Table 2 shows the information of sparse matrixes used in the performance evaluations in details. Five different sparse matrixes are used to compare and analyze the performance of SpMV. As shown in Table 2, the scarcity and the size of the sparse matrix are different from each other.

Table 3. The total execution time of each SpMV according to the number of CPU cores. (Unit: Second)

Matrix name	The number of iterations	The number of CPU cores used			
		1	8	16	32
road_usa	1	0.96	0.45	0.63	0.58
	500	224.52	35.93	21.81	12.62
	1000	448.70	68.99	43.01	24.68
audikw	1	0.38	0.24	0.14	0.09
	500	80.42	42.58	21.63	11.12
	1000	160.82	84.89	43.14	22.16
thermal	1	0.08	0.05	0.04	0.04
	500	19.04	3.16	1.68	0.79
	1000	38.06	6.27	3.32	1.53
msdoor	1	0.10	0.04	0.03	0.03
	500	21.83	4.58	2.93	1.49
	1000	43.53	9.12	5.82	2.95
twotone	1	0.01	0.01	0.01	0.00
	500	2.73	0.39	0.21	0.11
	1000	5.44	0.78	0.41	0.21

2.2 Evaluation Results

In this section, we will explain the results of comparing and analyzing the computation efficiency between multiple GPUs and multiple CPUs based cluster computing environments.

Tables 3 and 4 shows the total execution time results of each SpMV operation according to the multiple CPUs and multiple GPUs, respectively. The number of iterations means that the number of iteration of each SpMV operation. In short, there are 1000 loops in a SpMV operation in case of the number of iterations is 1000. As the number of iterations increases, the computation load increases. Generally, GPU is more efficient than CPU in terms of execution time in case of the larger computation loads.

Table 4. The total execution time of each SpMV according to the number of NVIDIA GeForce Titan Black. (Unit: Second)

Matrix name	The number of iterations	The number of GPUs used			
		1	2	3	4
road_usa	1	0.49	0.48	0.59	0.81
	500	3.53	2.00	1.67	1.59
	1000	6.64	3.56	2.72	2.40
audikw	1	0.21	0.22	0.36	0.52
	500	16.56	13.04	11.85	11.23
	1000	32.96	25.90	23.56	22.32
thermal	1	0.04	0.16	0.30	0.44
	500	0.51	0.39	0.48	0.58
	1000	0.95	0.62	0.64	0.71
msdoor	1	0.06	0.17	0.30	0.44
	500	3.13	1.88	1.73	1.45
	1000	6.21	3.73	3.17	2.47
twotone	1	0.02	0.28	0.42	0.59
	500	0.27	0.44	0.60	0.66
	1000	0.50	0.61	0.54	0.59

As shown in Tables 3 and 4, in most cases, as the number of GPUs and CPUs used increases, the execution time of SpMV operations decreases. Overall, the GPU shows better performance than the CPU in terms of the total execution time.

However, even if the number of GPUs used increases, the execution time of SpMV does not decrease in some cases of the GPU. In the case of GPU, the necessary data must be present in the global memory of the GPU in order to do computational operations. Therefore, this is due to the overhead that is occurred by memory copy operations between the main memory and global memory of GPU. In short, the memory copy overhead between main memory and global memory of GPUs is greater than the reduced computation time by using multiple GPUs.

Based on the evaluation results, we can figure out that multiple GPUs do not always show better performance than multiple CPUs. Therefore, it is necessary to selectively use the CPU and the GPU in consideration of the characteristic of the operation.

3 Conclusions and Future Works

Cluster computing environments are popularly used in many applications that are sensitive to the computation speed. Because the computing power of cluster computing environments is much higher than a typical single node based computing system. In a cluster computing environment, multiple CPUs or multiple GPUs can be used for computation. This makes it possible to efficiently accelerate various computational operations.

In this paper, we evaluated and analyzed the performance of SpMV by using multiple GPUs and multiple CPUs based cluster computing environment. According to the performance evaluation results, the performance of SpMV operation improved as the number of GPUs and CPUs used increases in common cases. In the case of multiple GPUs, there are some unusual cases that the performance of SpMV in terms of total execution time decreased as the number of GPUs increased. Because the memory copy overhead between main memory and global memory of GPUs is greater than the reduced computation time by using multiple GPUs.

Therefore, we have to consider both memory copy overhead and computation overhead when using multiple GPUs and multiple CPUs. In the future works, we will study a policy that efficiently allocates computational operations by considering the computational efficiency of GPU and CPU in terms of memory copy overhead, computation overhead, and the characteristics of matrixes.

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No.2017R1C1B5017476). The corresponding author is Jeong-Dong Kim.

References

1. Kim, B., Jung, H.: A case study of data transfer efficiency optimization for GPU- and infiniband-based clusters. In: the 2015 Conference on Research in Adaptive and Convergent Systems, pp. 247–250. ACM, New York (2015)
2. Awan, A.A., Hamidouche, K., Hashmi, J.M., Panda, D.K.: S-caffe: co-designing mpi runtimes and caffe for scalable deep learning on modern gpu clusters. In: ACM Sigplan Notices, vol. 52, no. 8, pp. 193–205 (2017)
3. Stone, J.E., Saam, J., Hardy, D.J., Vandivort, K. L., Hwu, W.-M.W., Schulten, K.: High performance computation and interactive display of molecular orbitals on GPUs and multicore CPUs. In: 2nd Workshop on General Purpose Processing on Graphics Processing Units, pp. 9–18. ACM, New York (2009)

4. Kim, B., Jung, J., Min, H., Heo, J., Jung, H.: Performance evaluations of multiple GPUs based on MPI environments. In: the 2017 Conference on Research in Adaptive and Convergent Systems, pp. 303–204. ACM, New York (2017)
5. Panda, D.K., Tomko, K., Schulz, K., Majumdar, A.: The MVAPICH project: evolution and sustainability of an open source production quality MPI library for HPC. In: Workshop on Sustainable Software for Science: Practice and Experiences, Held in Conjunction with Int'l Conference on Supercomputing (2012)

Remote Block Heater Controller

Samuel Garbo, Zhazhuli Pratama Nur Winaziz, Ermal Elbasani, Jae Sung Choi, and Hyun Lee(✉)

Division of Computer Science and Engineering, Sun Moon University, Asan, Chung-Nam, South Korea
samuel.garbo@gmail.com, zhazpnw22@gmail.com, ermal.elbasani@gmail.com, {jschoi,mahyun9l}@sunmoon.ac.kr

Abstract. In general, an engine block heater helps the engine start and reduces fuel consumption and engine wear. Block heaters are electric heating elements that warm up the engine via different methods depending on the specific design. It prevents the coolant from gelling or freezing and stops the oil from turning into tar. In regions with cold winter private and public parking lots provide an electrical outlet for the car's block heater. Typically, the electrical outlet's heating timer is set manually. In this paper, we propose an IoT element to the existing electrical outlets that enable the heating timer to be set up remotely with automated heating time.

Keywords: Block heater controller · IoT cloud service · Real time clock · Thingspeak · Remote controller

1 Introduction

In region with cold winters such as the northern United States, Canada, Russia and Scandinavia, block heaters are frequently used for supporting a vehicle control. Particularly, block heaters are often standard equipment in new vehicles. For example, electrical outlets are found in public or private parking lots especially in multi-story car parks in extremely cold climates. In this case, a block heater warms an engine to increase the chances that the engine will start as well as warm up the vehicle faster than it normally would in extremely cold weather [1]. Among lots of types, the most common type is an electric heating element in the cylinder block, connected through a power cord often routed through the vehicle's grille. The block heater may replace one of the engine's core plugs. In this fashion, the heater element is immersed in the engine's coolant, which then keeps most of the engine warm. This type of heater does not come with a pump. They may also be installed in line with one of the radiator or heater hoses. Some heaters pump and circulate the engine coolant while it heating, others only heat the still coolant in the reservoir. Block heaters or coolant heaters are also found on permanently installed systems using diesel engines to allow standby generator sets to take up load quickly in an emergency. Research by the Agricultural Engineering Department of the University of Saskatchewan [2] has shown that operating a block heater for longer than four hours prior to starting a vehicle is a waste of

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 416–421, 2020.
https://doi.org/10.1007/978-981-13-9341-9_72

energy. It was found that coolant temperature increased by almost 20 °C (36 °F) degrees in that period, regardless of the initial temperature (4 tests were run at ambient temperatures ranging from −11 to −29 °C or 12 to −20 °F; continued use of the heater for a further one or two, or more, hours achieved a mere 2 or 3 more degrees Celsius as conditions stabilized. Engine oil temperature was found to increase over these periods by just 5 °C (9.0 °F). There are alternatives to a block heater that offer some of the same benefits. These include heaters attached to the engine's oil pan, usually with magnets. Dipstick heaters can be installed in place of the engine's oil dipstick. Heated blankets are available for the entire engine area, as well. A timer can be used with any of these heaters, so that it does not have to be left on all the time. This can help lower the electrical costs of using a block heater [3].

Setting the heating timer [4] is usually still manual. This means that every time the vehicle needs to be warmed the owner needs to go outside to the electric outlet and set the timer manually. Therefore, this paper proposes an IoT implementation to enable remote control on the timer and automate the heating time to minimize waste of electricity. The electrical outlet is paired with an Arduino system [5] to enable remote control. When a time is set, the Arduino system calculates the heating time based on the temperature outside. The user sets the time through a smart phone application.

2 Proposal

The prototype consists of a smart phone application, IoT cloud service and IoT device. The smart phone application runs on an Android environment. Thingspeak's IoT cloud service is used for conveying data between the application and IoT device. In particular, the IoT device consists of an Arduino Uno microcontroller, a temperature sensor, a red led, a relay module, Ethernet shield and real time clock module as shown in Fig. 1.

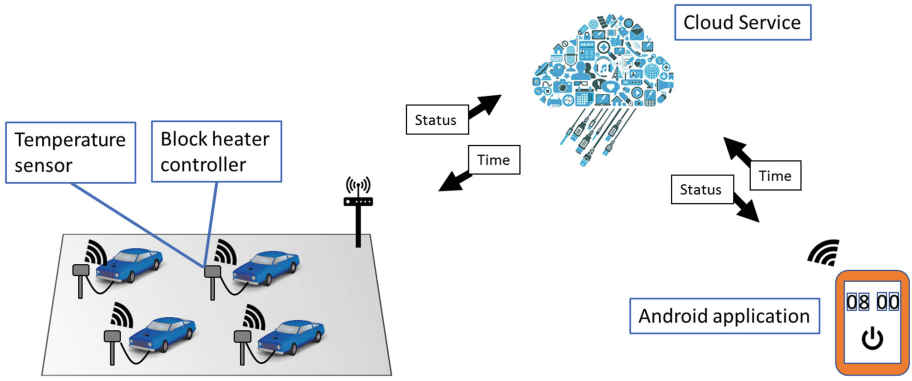


Fig. 1. Proposed remote block heater controller mechanism

3 Design and Implementation

The detailed implementation environment of the remote block heater controller proposed in this paper is shown in Table 1.

Table 1. Environment

	Environment	Description
H/W	Arduino Uno	Sensor and module controller
	TMP36	Temperature sensor
	DS3231	Real time clock module
	2-channel relay	Heating control
	Ethernet shield	Internet access
	5 mm Red LED	Heating status
S/W	MIT App Inventor 2	Android app development
	Arduino IDE 1.8.5	Arduino code

The configuration of the system is as follows. When the user has alarm settings and activates it from the installed Remote Block Heater Controller application on the smartphone, the alarm time data inputted by the user through the application on their smartphone will be transferred via the internet signal to the web data streaming service (Thingspeak) to be recorded. Then the alarm time data that Thingspeak has recorded will be forwarded to be sent to the Arduino system using an Ethernet (LAN) connection. Arduino will accept the data connection and immediately process it with the components that have been coupled and connected to each other with Arduino system. Components are a RTC (Real Time Clock) module, temperature sensor, relay module and an LED as a marker that the data connection has been successfully received by the Arduino system. Then the Arduino system will provide feedback to the data service center when a series of data transmission algorithms work, and Thingspeak will display a graph to show the data changes.

3.1 RTC (Real Time Clock) Module-DS3231 and Temperature Sensor-TMP36

This module [6] serves as a tool to synchronize alarm time that has been inputted by the user through the application on their smartphone with the real time in the world. So that the alarm time that has been inputted by the users can be detected in real by Arduino system.

In addition, temperature sensor(TMP36) is the most important part in the series of components that have been arranged into the Arduino system, because as has been explained that the main function of the Block Heater is as a warming system when the car in a state of very cold temperatures. Therefore, the sensor temperature [7] plays an important role to detect the temperature around the area so that the algorithm circuit in the system works based on the base time table of the Block Heater itself. For example,

we assume that if the temperature is below −10 °C, the heating timer operates for 2 h. If the temperature is from −10 to −5 °C, the heating timer operates for 1 h. If the temperature is over −5 °C, the heating timer operates for 30 min.

3.2 Relay Module 2-Channel

Module [8] works like a switch that acts to break and connect the flow of electricity on the Block Heater machine that has a voltage greater than the Arduino, so it functions like a voltage stabilizer when the device Arduino connected to the engine Block Heater. Not only that with 2-channel relay module users can choose 2 options that is using manual method to set the alarm time directly from Block Heater machine itself or activate it automatically through application installed on their smartphone. So, when one option is executed then the other option will not work because the electric current voltage is closed by the relay. Figure 2 shows that how relay module works with Block Heater machine.

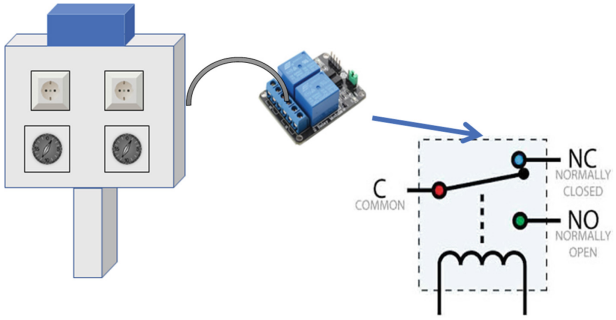


Fig. 2. How relay module works with Block Heater machine

3.3 Ethernet Shield (LAN) and 5 mm Red LED

Ethernet shield [9] that is integrated directly with the Arduino device serves as an internet connection by LAN cable between the Arduino device with server Thingspeak connecting the two of them to communicate with each other to exchange data. In addition, LED light is useful as a sign that heater has been activated. When the user starts setting alarm time from their smartphone application then turn on heater, then the LED light will be lights up. Contrariwise when there is no activity from the user or heater turns off after the timer ends then the LED lights will be lights down.

3.4 Implementation and Result

As a result, we make the user interface to control the Block Heater remotely. The user interface has 3 main functions. The first is the display of the current alarm time. The second is the control of the alarm time and the third is displaying the heating status. The application receives data from the IoT platform via polling.

The IoT platform gathers 3 types of data: hour, minute and heating status. The alarm time set by the user is stored separately in hours and minutes. When the alarm is turned off a negative value is stored. The heating status is constantly being updated by the IoT device. Figure 3(a) shows the example of a mobile application's UI and Fig. 3(b) shows the example of a Thingspeak IoT platform.

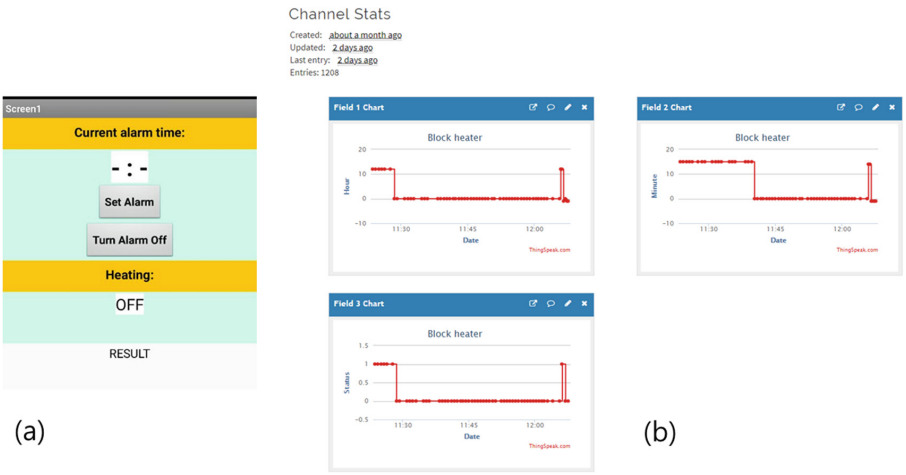


Fig. 3. (a) Mobile application's UI and (b) Thingspeak IoT platform

In addition, the IoT device connects to the IoT platform's REST API via polling. If an alarm time exists, the IoT device gets the current time and temperature and calculates the heating's starting time. If the current time is between the heating's start and end time the heating is turned on otherwise the heating is off. Figure 4(a) shows the example of the IoT device's algorithm and Fig. 4(b) shows the example of a hardware prototype.

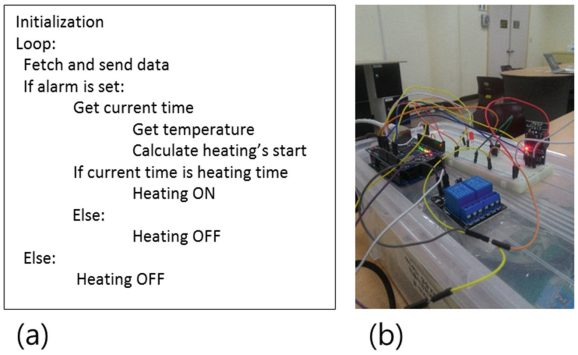


Fig. 4. (a) IoT device's algorithm and (b) Hardware prototype

4 Conclusion

This paper proposed an IoT device for controlling remotely the electrical outlet used for the car's block heater. With remote control available, setting up the timer is more flexible and comfortable to the user. The user doesn't need to be at home or nearby for setting or changing the timer. Also, with the automated heating time, the user doesn't need to estimate the heating time in advance and the block heater's electrical consumption becomes optimized.

Acknowledgments. This work is result of a study on the "Leader in Industry-university Cooperation +" Project, which is supported by the Korean Ministry of Education.

References

1. Block heater (2018). https://en.wikipedia.org/wiki/Block_heater
2. Finland: How to survive winter in Finland and enjoy it. https://finland.fi/life-society/how-to-survive-winter-in-finland-and-enjoy-it/
3. International Energy Agency (IEA) HEV: Finland – Charging infrastructure. http://www.ieahev.org/by-country/finland-charging-infrastructure/
4. Insideevs: Car Machine Warming Solutions in Winter. https://insideevs.com/finland-has-a-genius-charging-solution-for-evs-use-existing-block-heater-poles-video/
5. Arduino Uno: https://www.arduino.cc/en/Main/ArduinoBoardUno
6. How to Mechatronics: DS3231-real time clock tutorial: https://howtomechatronics.com/tutorials/arduino/arduino-ds3231-real-time-clock-tutorial/
7. Adafruit Website: Tmp36 temperature sensor. https://learn.adafruit.com/tmp36-temperature-sensor/overview
8. Instructables Website: Relay Module. http://www.instructables.com/id/Arduino-2-Channel-Relay/
9. Arduino Website: Arduino Ethernet Shield. https://www.arduino.cc/en/Guide/Arduino-EthernetShield
10. Blum, J.: Exploring Arduino, IT Cookbook (2014)

A Method of Intelligent Dynamic Monitoring for Real-Time Data Streaming Processing Systems

Onur Soyer(✉) and Kwan Young Park

Mobigen Co., Ltd., C-16th Floor, 128, Beobwon-ro, Songpa-gu, Seoul 05854, Republic of Korea
onrsyr@gmail.com, kypark@mobigen.com

Abstract. We introduce a method that dynamically calculating a threshold value in real-time using ML techniques to determine whether a large IT system maintains its targeted performance and whether it is operating normally without the intervention of the operator operating the system.

Keywords: Big data · Data streaming · Anomaly · System monitoring · Machine learning

1 Introduction

Real-time data streaming systems are composed of lots of complex nodes producing data itself or gathering it using sensors in a continuous fashion. System administrators have to monitor this complex systems continuously to prevent any kind of issue before it happens or at least fix the issue with a minimum downtime. However, doing that is not easy for such a complex systems with lots of dependencies.

Monitoring complex data streaming systems has been becoming more important topic each day due to the growing size of node as it mentioned in Moore's Law [1]. For some systems composed of lots of sensors such as telecommunication network, grid network require lots of man power to monitor every second. And each sensor/node generates a considerable amount of data.

In order to collect and process big-stream data in real time, system performance is of utmost importance. Therefore, the system should maintain its target performance and monitor whether it is operating normally. Some criteria are needed to determine if the system maintains the target performance. However, when the operator who operates the system designates and manages the reference value, it is necessary to adjust the threshold value one by one according to the change of the amount of input data and the system failure situation, but it is practically impossible to cope with the changing situation at any time.

To overcome this problem we have developed a method that dynamically calculating a threshold value for determining whether the monitoring system operates normally and maintains its targeted performance in real-time. By doing that we are able to determine whether a big stream data processing system works properly and are able to detect an anomaly in real-time.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 422–427, 2020.
https://doi.org/10.1007/978-981-13-9341-9_73

This threshold value is calculated by collecting the performance indicator value from a big stream data, gathering from nodes and sensors, processing system in real-time.

The method we developed is composed of machine learning algorithm described in detail following sections to adapt itself to changing environment and conditions without a human touch.

2 Similar Works

Several papers point out that using machine learning algorithms is inevitably make the system more dynamic and problem free. One of the those work made by D. Adams et al. found that using machine-learning algorithm could then be used to determine if changes in measured features are statistically significant in the context of wind speed variations [2] which helps to smart dynamic monitoring. Another research made by Iverson [3] take advantage of some of ML algorithms such as K-means clustering and density-based clustering to solve smart monitoring problem. Also Ciang et al. [4] points out that many of the damage detection techniques require algorithms to function, ranging from simple and direct to complex with learning capability. These algorithms are capable of processing the signals from the sensors for damage identification, damage localization, severity assessment, and can be further extended to the prediction of failure, estimation of the remaining service life and decision making to determine the action required for the damage detected.

As we can see the works done by other researchers tried to find the answer to similar question for specific systems they worked on.

3 Methodology

In this proposed method, we have used a statistical quality technique called Logistic Regression [5] used in SPC (Statistical Process Control) [6].

Statistical process control (SPC) is a method of managing processes by statistical methods to identify products that can pass quality standards by identifying and interpreting problems based on data and finding solutions and improvement measures.

Logistic regression is a statistical technique used to predict the likelihood of an event using a linear combination of independent variables. It is similar to linear regression in terms of describing dependent variables as a linear combination of independent variables, but there is a difference in that the dependent variable is categorical data, and when the input data is given, the result of the data is divided into specific categories.

This method specifies a real-time threshold adjustment methodology based on a type 1 error and a type 2 error in terms of performance SLO (Service level objective). The key idea of the proposed methodology is to model the SLO violation using Logistic Regression and then adjust the thresholds iteratively and retrospectively.

We use streaming sensor data to train the model in real-time. By doing this, our model can adapt itself to changing environment continuously and prevent damage before it happens.

This system of course might be given a pre-defined threshold value for the sake of calculation. Later then threshold value will be optimized with value's gathering form streaming sensor data.

4 Calculation

Let Γ (t) = {γ (0), γ (1), … ,? (t)}. A (t) is probabilistic. Taking a binary variable, assume a (t) = 0 for violation of SLO and a (t) = 1 value for violation. Y(t) is a stochastic binary variable. If y (t) = 1 and y (t) = 0 are violated, the performance indicator (data processing response time) The α and β values are defined by the target PPV (Positive Predictive Value) and NPV (Negative Predictive Value) values (see Table 1). The system operates for a sufficiently long period of time, (T + 1) of the next step should be calculated when converging to t → ∞, PPV → α, NPV → β, assuming that there is a continuous monotonicity.

Table 1. Threshold state compare to SLO

	SLO violation	SLO non-violation
Threshold violation	X	Y
Threshold non-violation	U	V

1. Error: 1-PPV;

$$\mathrm{PPV} = \frac{x}{x+y}. \tag{1}$$

2. Error: 1-NPV;

$$\mathrm{NPV} = \frac{v}{u+v} \tag{2}$$

In regression modeling, Y(t) is used as a response variable and Γ(t) and A(t) are specified as explanatory variables. The regression equation is then obtained by using PPV and NPV for γ(t + 1). Therefore, Γ(t) is also used as an explanatory variable in the regression model and as an estimate of the next step in the regression model.

The reason for using Γ(t) and A(t) is as follows. First, because Γ(t) is a control variable of Y(t) directly. If the threshold is taken too large, the threshold violation event will not occur and Y(t) will have mostly zero values. This increases the rate of the second kind error. On the other hand, if the threshold is taken very small, many threshold violations will occur and Y(t) will have a value of 1 and the error rate of the first kind will increase. A(t) is necessary to obtain the appropriate values of PPV and NPV.

Since Y(t) is a binary variable, the general linear regression equation $y = c + \vec{b} * \vec{x}$ cannot be used. The first reason is that if $\vec{x}$ is infinitely large, the y value exceeds 1. The second is that the general linear regression is based on the assumption that the significance test of the regression coefficient follows the normal distribution of the residual $\in = \acute{y} - y$. y has 0 and 1 values, so it does not match. Finally, A(t) is a categorical variable and Γ(t) is a continuous variable. For this reason, logistic regression is used as a logit function.

$$p(\vec{x}) = P\left(Y = 1 \middle| \vec{X} = \vec{x}\right) = 1 - P\left(Y = 0 \middle| \vec{X} = \vec{x}\right) \tag{3}$$

We obtain $\text{logit}(p(\vec{x})) = \ln \frac{p(\vec{x})}{1-p(\vec{x})}$ through logit transformation. Therefore, a general linear regression equation can be made using the logarithm value as a response variable.

$$\text{logit}(p(\vec{x})) = c + \vec{b} * \vec{x} \tag{4}$$

The regression coefficients c and $\vec{b}$ are obtained using the Maximal Likelihood Estimator.

$$p(\vec{x}) = \frac{1}{1 + \exp\left(-c - \vec{b} * \vec{x}\right)} \tag{5}$$

Say $x_1 \equiv a(t)$, $x_2 \equiv \gamma(t)$ and then the following equation is obtained.

$$p(t+1) = \frac{1}{1 + \exp(-c - b_1 * a(t+1) - b_2 * \gamma(t+1))} \tag{6}$$

a(t + 1) has a value of 0 or 1, and when the SLO is not violated;

$$p(t+1) = \frac{1}{1 + \exp(-c - b_2 * \gamma(t+1))} \tag{7}$$

If the SLO is violated;

$$p(t+1) = \frac{1}{1 + \exp(-c - b_1 - b_2 * \gamma(t+1))} \tag{8}$$

(8) As the propagation converges from $t \to \infty$ to a true positive alarm, the probability value of (8) can be expressed as PPV. (7) The reproduction probability value 1-p converges to NPV. Therefore, if PPV and NPV are expressed as α and β, the following equation can be obtained.

$$\gamma(t+1) = \frac{ln\frac{\alpha}{1-\alpha} - ln\frac{\beta}{1-\beta} - 2 * c - b_1}{2 * b_2} \tag{9}$$

(9) Implementation of the automatic threshold setting algorithm using pseudo code using reproduction is as follows.

```
ATS(α,β){
Specify the values of α and β //ex) α=0.95, β=0.9
          Specify the initial value of γ (0)
          while(TRUE){
            Given SLO at time t, do{
            Boolean v <- isViolated(SLO)
              if((v ∧ (μ(t) < γ(t))) ∨ (~v ∧ (μ(t) > γ(t)))
               if((current PPV < α) ∨ (current NPV< β))
            {
               fitting logistic regression
               Calculate γ (t + 1) using equation (9)
                }
              Else
               γ(t+1)<- γ(t)
              }//end:do
            }//end:while
          }end:ATS
```

5 Conclusion

When monitoring a large-scale IT system, it is practically impossible to set and manage an appropriate threshold value for each system due to a large number of devices to be managed.

When our model is applied to a monitoring system of a large-scale IT system, more efficient operation of the system is possible by injecting less manpower and time.

In addition, it can dynamically respond to changing operating environments and adapt itself to forecast the future, so that it is possible to grasp and prepare for the timing of capacity additions of IT systems in advance.

Acknowledgments. This work was supported by the Institute for Information & Communications Technology Promotion (IITP) with a grant funded by the Korean government (MSIT). (No. 2016-0-00193, IoT Security Vulnerabilities Search, Sharing, and Testing Technology Development).

References

1. Moore's Law. https://en.wikipedia.org/wiki/Moore's_law

2. Ciang, C.C., Lee, J.-R., Bang, H.J.: Structural health monitoring for a wind turbine system: a review of damage detection method. Meas. Sci. Technol. **19**, 122001 (2008). https://doi.org/10.1088/0957-0233/19/12/122001
3. Iverson, D.: Inductive system health monitoring, pp. 605–611 (2004)
4. Zhong, J., Yang, Z., Wong, S.F.: Machine condition monitoring and fault diagnosis based on support vector machine. In: 2010 IEEE International Conference on Industrial Engineering and Engineering Management, Macao, pp. 2228–2233 (2010)
5. Logistic Regression. https://en.wikipedia.org/wiki/Logistic_regression
6. Statistical Process Control (SPC). https://en.wikipedia.org/wiki/Statistical_process_control

A UAV Path Planning Method Using Polynomial Regression for Remote Sensor Data Collection

Kwang Min Koo[1], Kyung Rak Lee[2], Sung Ryung Cho[1], and Inwhee Joe[1(✉)]

[1] Department of Computer and Software, Hanyang University, Seoul, Korea
{codingple, kyougt, iwjoe}@hanyang.ac.kr
[2] Unmanned Vehicle Systems Research Group, Electronics and Telecommunications Research Institute, Gajeong-ro, Daejeon 34129, Korea
krlee@etri.re.kr

Abstract. The use of Unmanned Aerial Vehicle (UAV) has been increasingly diverse. In Wireless Sensor Networks (WSNs), UAVs are used for sensor data collection. However, the path planning for the UAV is not practical for a large number of sensors, and is not effective for the battery consumption. In this paper, we propose the UAV path planning for data collection in WSNs using Polynomial Regression. With this algorithm, the UAV is able to save the battery and the mission time.

Keywords: UAV · Path planning · Sensor data collection · Machine learning

1 Introduction

Gradually, the use of Unmanned Aerial Vehicle (UAV) has been amplified at the versatile viewpoints. Likewise, UAVs assist to efficiently collect data. They are able to explore where people cannot or have a difficulty to approach and to perform a cumbersome data collection. This enables data acquisition in the inaccessible, disastrous, and extensive area. To gather environmental data, unmanned vehicles have been used at sea, in the air, in dessert, even on other planets [1]. Moreover, the quick image-processing method of UAV was researched for remote sensing in earthquake disaster area [2]. As an alternative to traditional multi-hop communication schemes, sensor data collection in a huge area with UAV is more beneficial in terms of the sensor battery lifetime [3].

Accordingly, researches to utilize UAVs in Wireless Sensor Networks (WSNs) have been conducted for sensor data collection as well. In WSNs, it is a significant factor for the effective data acquisition to design the adequate path, in that this influences the availability of UAV in terms of its battery and the efficiency of data exchange associated with the distance between UAV and sensor node. Nevertheless, the path planning for UAV collecting sensor data is not sufficiently studied. The previous research is not

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 428–433, 2020.
https://doi.org/10.1007/978-981-13-9341-9_74

practical due to the number of nodes to obtain data. Besides, the traditional path planning is not a proper solution for data collection to optimize battery efficiency.

In this paper, we propose a UAV path planning method for sensor data collection in WSNs using polynomial regression with GPS coordinates of sensor nodes. The regression based on latitude and longitude of nodes generates a continuous curved path across the sensor area, which is formed an equation. The length of this path is shorter than the general path, which means the battery efficiency and mission time of UAV is more practical.

2 Related Works

In WSNs, researches to utilize UAVs have been conducted for sensor data acquisition. However, the path planning for UAV collecting sensor data in WSNs has not been sufficiently developed. Some study configures the predetermined path with simple patterns. The paper [4] assumes the path is predefined linear. Other papers also define the mobility pattern as circular [5], or Square, Angle, Tractor Mobility [6]. These papers mentioned above focus on other purposes, e.g., the communication effectiveness, fairness, et al.

In addition, the traditional path planning for UAV is not a proper method. One of the general methods is Genetic Algorithm, which randomly combines the partial factor of feasible cases to find the optimal output [7–9]. The map UAVs explore is divided as a grid, and the factor is a part of the sequential cells of the grid. The algorithm makes the optimal path with the iterative generation. However, it is not regarded to collect sensor data, which means the entire sensors become the waypoints of the output. This trajectory is able to waste the battery and operation time of UAV to collect data. Another path planning method is Voronoi Diagram [10, 11]. This algorithm separates the area points are deployed on using the normal lines at the spot where the distance is equal between two the points. The purpose of these planning methods is collision avoidance, search and rescue operations, physical security from threat sources, e.g.

The paper [12] proposes the optimizing solution of UAV's trajectory and wake-up schedule of Sensor nodes for saving energy. This work uses convex optimization approach with the various parameters including data size, transmission power, etc. As a result, the path is optimized in that it does not definitely visit every node for battery efficiency. However, the complicated optimization enforces the path to deal with the few number of nodes. It needs more practical and simple solutions to handle a number of sensors.

3 Path Planning Using Polynomial Regression

Regression is one of machine learning algorithm used in data prediction. In terms of the data fitting process called training, which minimizes the loss function, it is an optimization method as well. Using this method, the path for the battery efficiency is calculated.

3.1 Regression with Coordinates of Sensor Nodes

The input of the proposed method is GPS coordinates of each node. Since the coordinates consist of longitude and latitude, the derived equation of the proposed path planning is represented by the relationship between them, which enables this result to be used for 2-dimensional path. The following Eq. (1) is the hypothesis the regression assumes for training, and the Eq. (2) describes the output path as 2-dimensional GPS containing the Eq. (1).

$$H(x) = \sum_{t=0}^{n} \theta_n x^{n-t} \tag{1}$$

$$Path = (x, H(x)) \tag{2}$$

The hypothesis is an n^{th} degree polynomial equation where x is the longitude of the input. The weight for each x, denoted θ, is changed to fit the input in the training process. In other words, the training modifies θ in order that $H(x)$ becomes approximate to the corresponding latitude of each longitude x. As a result, the path for the UAV is able to describe as the Eq. (2) where $H(x)$ is trained with the data and x is limited within the range between the maximum and the minimum longitude of the sensors coordinates.

3.2 Modified Loss Function for Transmission Range

Since the input of the regression is the GPS coordinates, the loss function of the proposed regression is defined as the distance from each sensor to the path. In common, Mean Squared Error (MSE) is used for loss function. The following Eq. (3) describes MSE.

$$L(x) = \sum_{t=0}^{n} \frac{1}{n} (y_t - H(x_t))^2 \tag{3}$$

The predicted latitude of the t^{th} input data by the training is $H(x_t)$, and the real latitude is denoted by y_t. The θ in (1) is modified to minimize $L(x)$ value. However, this process possibly causes that there is the data transmission range of the sensors excluding the trajectory. Since the value of $L(x)$ more reduces when the path moves into the area where the sensors are concentrated, the coverage of the opposite side may exclude the flying UAV. For this reason, the loss function is modified as the following:

$$L'(x) = \sum_{t=0}^{n} \frac{1}{n} \omega_t (y_t - H(x_t))^2 \tag{4}$$

where ω is the weights for each distance from the path to the sensors. The weights for the sensor excluding the trajectory will be set relatively bigger than ones including. In addition, to prevent $L'(x)$ from diverging to infinity, the method of decreasing ω is adopted.

3.3 The Whole Path Planning Process

First of all, the all of θ in (1) is initialized randomly, and all of ω in (4) is initialized with 1. In the first loop for e_1, which denotes the specified number of iteration in advance, the regression is proceeded for calculating the sufficient path based on the distance from the path to the nodes. After the loop, the second loop considers the transmission range. In the second loop for e_2, which denotes the specified number of iteration in advance, ω for the range including the trajectory decreases gradually. In this loop, the sensors that have the coverage excluding the trajectory are identified each specified i times within e_2. At that time, the d multiplies to ω of the identified sensor, where $0 < d < 1$.

4 Simulation Results

A hundred of input coordinates is randomly sampled within the range between −3 and 3, assumed that they are latitude and longitude. In our simulation, the e_1 is 1000, e_2 is 2000, i is 50, and d is 0.8.

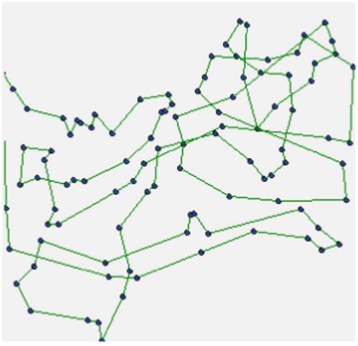


Fig. 1. Actual path from genetic algorithm

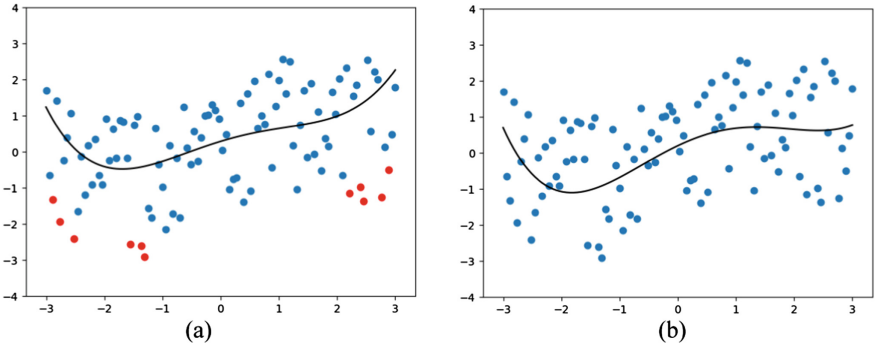


Fig. 2. Actual path from polynomial regression for first loop (a) versus whole loop (b)

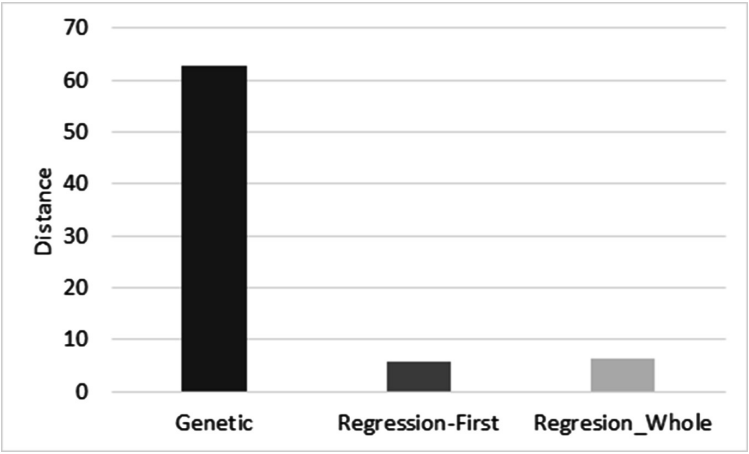


Fig. 3. Travel distance

Figure 1 visualizes the path planning using the traditional method, Traveling Salesman Problem with Genetic Algorithm. The (a) of Fig. 2 illustrates the first loop of the proposed method, and (b) shows the end of the algorithm. In (a) of Fig. 2, the red points are the sensors that have the range excluding the path. As the (b) of Fig. 2, the last loop allows the all coverage to cover the path. In the Fig. 3, the distance of Fig. 1 is about 62.66, (a) of Fig. 2 is 5.82, and (b) of Fig. 2 is 6.22.

5 Conclusion

We proposed the path planning for UAV carrying out data collection in WSNs, which derives the path of which the distance declines into 10% against the traditional path planning. Although the distance of (a) is shorter than (b), the path of (b) is able to pass all transmission range of the sensors for data transfer. For the future work, we will improve the loss function of the path planning to modify the curve faster and accurate.

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science, ICT & Future Planning) (NO. 2016R1A2B4013118).

References

1. de Sousa, J.B., Gonçalves, A.: Unmanned vehicles for environmental data collection. Clean Technol. Environ. Policy **13**(2), 369–380 (2011)
2. Li, C.-C., Zhang, G.-S., Lei, T.-J., Gong, A.-D.: Quick image-processing method of UAV without control points data in earthquake disaster area. Trans. Nonferrous Met. Soc. China **21**(Supplement 3), 523–528 (2011)

3. Wei, P., Gu, Q., Sun, D.: Wireless sensor network data collection by connected cooperative UAVs. In: 2013 American Control Conference, pp. 5911–5916, Washington, D. C. (2013)
4. Ma, X., Kacimi, R., Dhaou, R.: Fairness-aware UAV-assisted data collection in mobile wireless sensor networks. In: 2016 International Wireless Communications and Mobile Computing Conference (IWCMC), pp. 995–1001, Paphos (2016)
5. Abdulla, A.E.A.A., Fadlullah, Z.M., Nishiyama, H., Kato, N., Ono, F., Miura, R.: An optimal data collection technique for improved utility in UAS-aided networks. In: IEEE INFOCOM 2014, pp. 736–744, Toronto (2014)
6. S. Rashed, M. Soyturk: Effects of UAV mobility patterns on data collection in wireless sensor networks. In: 2015 IEEE International Conference on Communication, Networks and Satellite (COMNESTAT), pp. 74–79, Bandung (2015)
7. Li, J., Huang, Y., Xu, Z., Wang, J., Chen, M.: Path planning of UAV based on hierarchical genetic algorithm with optimized search region. In: 2017 13th IEEE International Conference on Control & Automation (ICCA), pp. 1033–1038, Ohrid (2017)
8. San Juan, V., Santos, M., Andújar, J.M.: Intelligent UAV map generation and discrete path planning for search and rescue operations. Hindawi Complex. **2018**(6879419) (2018)
9. Yang, Q., Yoo, S.J.: Optimal UAV path planning: sensing data acquisition over IoT sensor networks using multi-objective bio-inspired algorithms. IEEE Access **6**, 13671–13684 (2018)
10. Chen, X., Chen, X.: The UAV dynamic path planning algorithm research based on Voronoi diagram. In: The 26th Chinese Control and Decision Conference (2014 CCDC), pp. 1069–1071, Changsha (2014)
11. Chen, X., Li, G.Y., Chen, X.M.: Path planning and cooperative control for multiple UAVs based on consistency theory and Voronoi diagram. In: 2017 29th Chinese Control and Decision Conference (CCDC), pp. 881–886, Chongqing (2017)
12. Zhan, C., Zeng, Y., Zhang, R.: Energy-efficient data collection in UAV enabled wireless sensor network. EEE Wirel. Commun. Lett. **7**(3), 328–331 (2018)

A Portable Finger Language Translator Based on Deep Learning with Leap Motion

Dahye Jin, Yumi Lee, Ermal Elbasani, Jae Sung Choi, and Hyun Lee(✉)

Division of Computer Science and Engineering, Sun Moon University, Asan, Chung-Nam, Republic of Korea
{jd_l04, dldbalqk}@naver.com, ermal.elbasani@gmail.com, jschoi@gmail.com, mahyun91@gmail.com

Abstract. Hearing impaired people are less able to communicate than ordinary people, and experience many inconveniences in everyday life, hospitals, and government offices. To solve this problem, various geophysical translators and sign language translators are being developed. However, most of the wearable translator has the disadvantage that the accuracy of output data is low. Thus, we try to solve this problem by using leap motion with smart phone, which are hand motion devices, instead of wearable translator. Particularly, we tried to increase the recognition rate of geographical interpreters by applying a deep learning model with multiple perceptron. Experimental results show that the average recognition rate is almost 94.9% separately.

Keywords: Portable language translator · Geophysical translator · Leap motion · Deep learning · Smart phone

1 Introduction

In the case of the Republic of Korea, the number of people with hearing and speech impairments is about 300,000 and it is steadily increasing every year. In general, hearing impaired people are not easy to communicate with the general public. There are many inconveniences in daily life. To solve this problem, the nation operates a sign language interpretation center. However, the number of sign language interpreters nationwide is 455, which is insufficient supply compared to people who need service [4]. In addition, geographical translator for the hearing impaired is not popularly sold compared to necessity, and technical research is also small. For example, existing geographical translator translates geography into English, so it is difficult to use those who do not know English well [5–7].

In order to solve this problem, a low-cost, portable geared interpreter that can be carried easily is needed. In this study, we intend to develop a portable geophysical translator that recognizes and translates geophysical motion by linking lip motion and smartphone, which are hand motion recognition devices. In particular, we will try to increase the recognition rate of geographical interpreters by applying a deep learning model using multiple perceptron. The rest of paper is organized as follows. Chapter 2

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 434–439, 2020.
https://doi.org/10.1007/978-981-13-9341-9_75

introduces related work. Chapter 3 shows the proposed system model. In chapter 4, we make the experiment result and conclude the paper.

2 Related Work

There are many kinds of motion recognition devices that can grasp the movement of the hand such as a camera type, a ring type, and a bio signal type. Among them, Nod is a ring type motion recognition device. The Nod [1] uses a proprietary technology called 'Micro Gestures' to capture the movement of the fingers to the minute vibrations. It can work with various smart devices such as Android, iOS, Smart TV, Google Glass and so on. However, in the case of the Nod, the movement of the finger is well caught, but it is disadvantageous to be worn on all five fingers for translation. Figure 1(a) shows an example of Nod Ring.

The MYO armband [2] is an arm band type gesture control wearable device worn on the arm. In the case of MYO, if you move your arms and hands after wearing them on your wrist, you will recognize the movement of the muscles and transmit the data through Bluetooth communication. It can be used to control various digital devices such as Windows, Mac OS, Android and iOS. It is easy to wear on the arm, but it needs to be calibrated individually because the muscles move according to the user and the signal changes depending on the position to wear. There is a disadvantage that the position of the band can be changed while sweating or moving. Figure 1(b) shows an example of MYO armband.

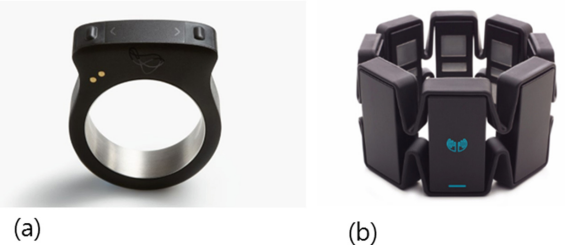


Fig. 1. (a) Nod Ring and (b) MYO armband

In addition, we applied the multi-perceptron [3] that is a neural network with one or more intermediate layers between the input and output layers. It has a hierarchical structure. The intermediate layer between the input layer and the output layer is called a hidden layer. The network is connected in the direction of the input layer, the hidden layer and the output layer, and is a feedforward network in which there is no direct connection from the output layer to the input layer in each layer.

As the number of layers increases, the characteristics of the crystal zone formed by the perceptron are further enhanced in the multilayer perceptron. In other words, it is possible to perform a backpropagation learning algorithm that can learn the hidden

layer rather because the analysis is a little complicated but differentiation is possible. Thus, most multi-layer perceptron can be learned using backpropagation learning algorithms. For example, when input data is presented to each unit of the input layer, the signal is converted in each unit, transferred to the intermediate layer, and finally output to the output layer. This output value is compared with a desired output value and the connection strength is adjusted in a direction to reduce the difference.

3 System Model and Implementation

The proposed system model of the portable geographical translator is divided into input part, processing part and output part according to each function as shown in Fig. 2. The input part connects the lip motion to the smartphone, receives the finger information of the user, and transmits the information to the server. In the processing part, the deep learning model which matches the transmitted data is found in the deep learning model, and the translated geographical data is transmitted to the user's smartphone. The output part outputs the transmitted image to the user's smartphone.

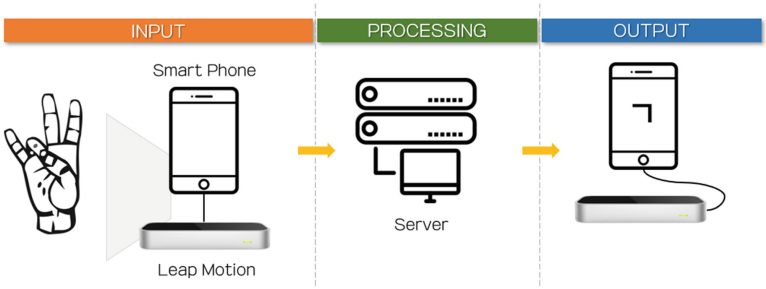


Fig. 2. The proposed system model

The input part receives the information of the geographical information that the user wants to translate, sends it to the server, and executes the application to obtain the geographical information. Particularly, the application determines whether or not the lip motion is connected. If the lip motion is connected, the application fetches the motion data of the hand in the recognition range, and transmits the fetched data to the server through the socket communication. Figure 3(a) shows the hardware structure of smartphone and lip motion applied in this study. Figure 3(b) shows the human finger structure used in the lip motion sensor [8–10].

In processing part, the distinction between the right hand and the left hand and the direction of each finger must be distinguished in order to discriminate the jig from the lip motion sensor. Then, the data received from the processing part and the deep learning model is searched for coincidence. If there is a match, the smartphone communicates through socket communication. However, unlike the PC version of the lip motion sensor that supports the direction of the hand, the direction of the finger, and the

direction of the finger, the Android version does not support the direction of the finger. Therefore, in this study, the direction of the hand was added as an attribute for judging the geomorphology as shown in Fig. 3(c). The directional values of lip motion support pitch, roll, and yaw values. Finally, the attributes of the model are distinguished from the right hand and the left hand, and the three directions of the five fingers are distinguished. Therefore, the model has 19 attributes.

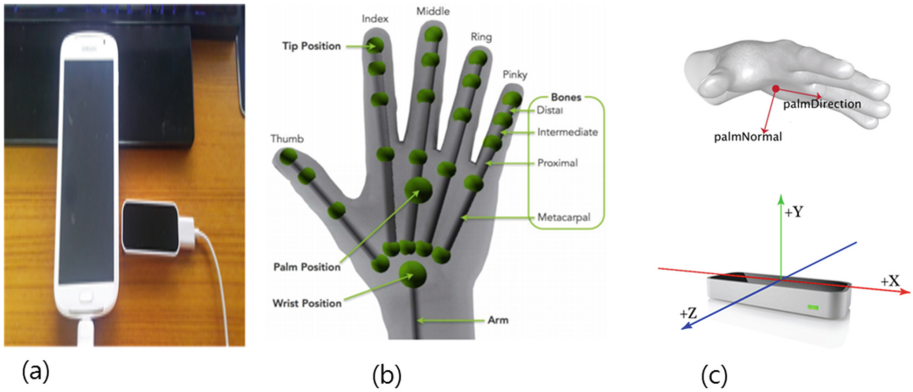


Fig. 3. (a) Smart phone and lip motion sensor, (b) Finger measure position, and (c) The added hand direction

In addition, datasets were created using Android and php, apache, and mysql. The total numbers of data for Deep Learning was 88190, and there were prepared a total of 29 consonants (14 pieces) and vowels (19 pieces) in each 1500–4000 pieces. Figure 4 shows what you are learning with the dataset. Moreover, the output part outputs the received data from the server to the smartphone screen.

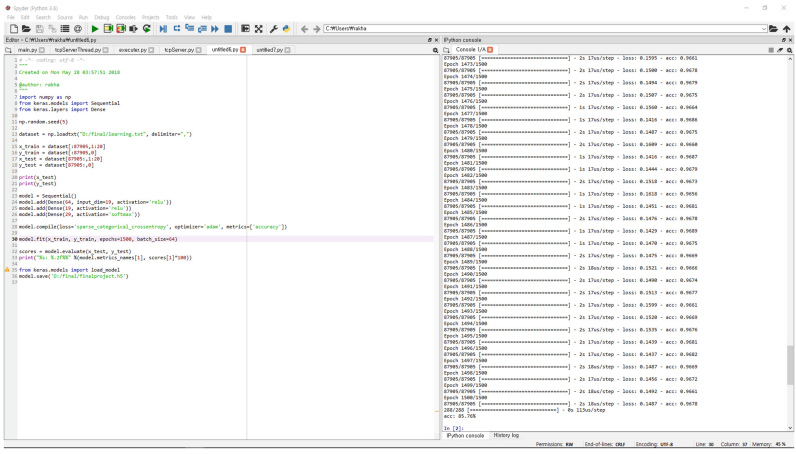
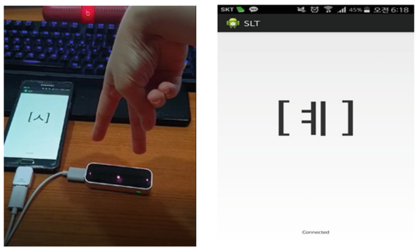


Fig. 4. An example of the learning process with the dataset

4 Experiments and Analysis

In this study, the recognition rate of the geographical translator was experimented. The lip motion connected to the smartphone was placed on the floor and measured 10 times in 10 consecutive consonants. Figure 5(a) shows the use of the geophysical translator according to the system model, and Fig. 5(b) shows the screen of the application that translates the geophysical translation. In addition, Table 1 shows the recognition rate of the geophysical translator.



(a) (b)

Fig. 5. (a) The use of the geophysical translator and (b) An example of the application result

According to the experiment result, the recognition rate of Korean consonants "ㄷ" is significantly lower than other consonants and vowels. This is because the hand direction of consonant "ㄷ" and "ㄹ" and the direction of the finger are very similar. This can be solved by increasing the recognition rate by further refining the attributes of the deep learning model if the lip motion Android support extends to the finger node.

Table 1. The experiment of the recognition rate of the geophysical translator

		Rate		Rate		Rate		Rate		Rate
Cons.	ㄱ	98%	ㄴ	97%	ㄷ	17%	ㄹ	92%	ㅁ	100%
	ㅂ	95%	ㅅ	100%	ㅇ	92%	ㅈ	99%	ㅊ	99%
	ㅋ	97%	ㅌ	99%	ㅍ	96%	ㅎ	95%		
Vowel	ㅏ	100%	ㅑ	99%	ㅓ	100%	ㅕ	98%	ㅗ	98%
	ㅛ	100%	ㅜ	100%	ㅠ	99%	ㅡ	95%	ㅣ	99%
	ㅐ	99%	ㅒ	99%	ㅖ	97%	ㅚ	95%	ㅟ	99%

5 Conclusion

The purpose of this study is to develop a portable geophysical translator to solve the inconvenience of the hearing impaired in everyday life. At present, the translator is commercialized in foreign countries, and the company is the main customer, and it is

necessary to purchase and use heavy tablets at high prices. However, the portable geophysical translator developed in this study, only needs to connect cheap and light lip motion to the smartphone. Most of them were connected to Raspberry PI or PC, but the geographical translator of this study is a big change because it is connected to smart phone and high portability. The translator of this study can translate 14 consonants and 19 vowels, and the average recognition rate is almost 94.9%. This recognition rate can be higher if lip motion's Android support is extended.

In the future research direction, we will develop a voice translation function to add voice to the text expressed in text, and to translate not only text but also sign language through lip motion gesture recognition function.

Acknowledgments. This work is result of a study on the "Leader in Industry-University Cooperation +" Project, which is supported by the Korean Ministry of Education.

References

1. Nod: The resource is available at https://nod.com
2. Myo-Gesture Control Armband: The resource is available at https://www.myo.com/
3. Kim, S., Kim, J., Ahn, S., Koo, B., Kim, Y.: An armband-type finger language recognition system based on ensemble artificial neural network. J. Korean Soc. Precis. Eng. **35**(1), 13–18 (2018)
4. Lee, J.W.: A study on the improvement of communication accessibility for the hearing impaired. In: Ministry of Health and Welfare Research Report (December 2013)
5. Jeong, P.S., Cho, Y.H.: Design and implementation of finger language translation system using raspberry pi and leap motion. J. Korea Inst. Inf. Commun. Eng. **19**(9), 2006–2013 (2015)
6. Jo, J.H., Kim, Y.R., Kim, H.J., Lee, S.K., Ro, K.H.: A research on a portable sign language translator with a smart device and a leap motion. Korea Intell. Inf. Syst. Soc. 772–775 (2016)
7. Kim, J.Y., Lee, J.G., Kim, D.J., Suh, Y.J.: Deep learning-based sign language recognition method using WiFi channel state information pattern. In: Proceedings of Symposium of the Korean Institute of Communication and Information Sciences, pp. 1435–1436 (2018)
8. Nowicki, M., Pilarczyk, O., Wasikowski, J., Zjawin, K.: Gesture recognition library for leap motion controller. Bachelor's Thesis in Poznan University of Technology (2014)
9. McCartney, R., Yuan, J., Bischof, H.-P.: Gesture recognition with the leap motion controller. In: International Conference on IP, Computer Vision and Pattern Recognition (IPCV 2015), pp. 3–9 (2015)
10. Hisham, B., Hamouda, A.: Arabic static and dynamic gestures recognition using leap motion. J. Comput. Sci. **13**(8), 337–354 (2017)

Research on MIDI Audio Watermarking Algorithm Based on Time Dynamics

Zhaolong Liu, Liang Chen, and De Li(✉)

Department of Computer Science, Yanbian University, Yanji, China
lll9160747@qq.com, 371482788@qq.com,
leaderl223@ybu.edu.cn

Abstract. The MIDI file is made up of hexadecimal digital control instructions, and it also has highly vulnerable. According to the characteristics of image watermarking technology and MIDI file, this paper improves the algorithm of embedding information in a single parameter by LSB method, and proposes a method of embedding digital watermarking in both time and strength parameters. This method can increase the embedding capacity and reduce the loss of sound quality and you can do blind extraction. At the same time, this paper analyzes the quality problem of MIDI audio embedded in watermark, and then puts forward some objective evaluation methods of MIDI audio.

Keywords: MIDI · Instruction parameter · Least significant bit · Estimation of quality

1 Introduction

Digital watermarking is a means to protected copyright. it is an important branch of steganography. Steganography is a way to hide secret information in an unrelated carrier, the secret information is known only to the person who made the contract, and no one else can know the hidden information even if they get the carrier. The core of steganography is cryptography. The digital watermarking technology is to embed some signs (i.e. digital watermark) directly into the digital carrier (including multimedia, document, software, etc.) or indirectly (modifying the structure of the specific area), without affecting the use value of the original carrier, and it is not easy to be explored and modified again. But it can be identified and identified by the producer. Through these information hidden in the carrier, can be achieved to identify the content creator, the buyer, the transmission of secret information or whether the carrier is tampered with or not [1].

MIDI music is a kind of music format that emerged twenty or thirty years ago, it is a combination of modern electronics, computer technology and some electronic audio processing equipment with music. MIDI audio is not a sound file, MIDI audio is a set of instruction notices inform equipment how to sound. According to the characteristics and vulnerability of MIDI file, this paper improves the algorithm of embedding watermarks only in a single parameter, which can increase the embedding capacity and reduce the loss of sound quality. This paper also made the quality evaluation after

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 440–445, 2020.
https://doi.org/10.1007/978-981-13-9341-9_76

embed watermarked, combined with the MOS subjective evaluation method of audio, several objective evaluation methods of MIDI audio quality are proposed.

2 MIDI Audio Watermarking Algorithm

Because of the particularity of the MIDI audio format, there are four kinds of hidden algorithms: the LSB minimum bit mode, the writing cannot write adjacent same instructions, the system special message mode, and the insensitive way of using MIDI to command order. The MIDI time parameters byte than the intensity parameters byte more of complex. Time difference refers to the time number of the previous event, and its unit is tick (the minimum time unit of MIDI) [2].

The least significant bit method is the simplest way of data embedding. The LSB algorithm treats arbitrary secret data as a string of binary streams, and each sampling data of the audio file is also represented by binary numbers. In this way, the most unimportant position of each sample value (in most cases is the lowest) can be replaced by binary data representing the secret data so as to achieve the purpose of encoding secret data in the audio number.

Increase the embedding capacity is bound to loss of sound quality, in does not affect the sound quality as far as possible under the premise of watermarking can be embedded into audio redundant information in order to improve the embedding capacity, in order to improve the efficiency of embedded can hide the secret information by minor changes to the original data, and can reduce the changes to the statistical characteristics of the original data, enhance the ability to resist steganalysis analysis. There are many ways to improve the embedding efficiency: sparse coding is proposed in document [3], and run coding is proposed in document [4]. In this paper, the algorithm uses LSB algorithm to embed watermark in time and intensity parameter to achieve the purpose of expanding the embedded capacity. The blind extraction algorithm is embedded in the inverse process. Embedding algorithm is as follows:

Suppose the watermark information is represented as 0 and 1bit strings, which is written as wm_i, the length of the watermark information is represented by *watermarked*. The length of the carrier MIDI file is $length[n]$, Fig. 1 is the basic block diagram of MIDI audio embedding algorithm.

Step1: i = 0, t = 0;
Step2: if(wa[i]&0xf0) = 0 × 90, k_{time}[i] = wa(i-1), $h_{strength}$[i] = wa(i + 2);
Step3: Count the number of occurrences of a value in the array k[i], called the number with the most number of occurrences is max1, the number with the second largest number is max2;
Step4: execute average = (max1 + max2)/2, the threshold value of blind extraction is recorded;
Step5: if(average)mod2 = 0 is true, return to Step6; else average = average + 1, next return to Step6;
Step6: if(max1 > max2), q = max1, else q = max2;
Step7: if(q)mod2 = 0 is true, q = q+2, else q = q+1, next return to Step8;
Step8:*if* $(k_i \geq q)$ and t $\leq$ watermarked, the embedding method is as follows:

D = k(i)-average
wa(h(i)) = average + bitset(d, 1, wm(1,t))
T = t+1, i = i+1;
Step9: if t > 0&t < watermarked and if wa(i + 2) > 50, wa(i + 2) = h(i)^wm(j), I = i+3, t = t+1; (notice wa(i + 2) > 50);

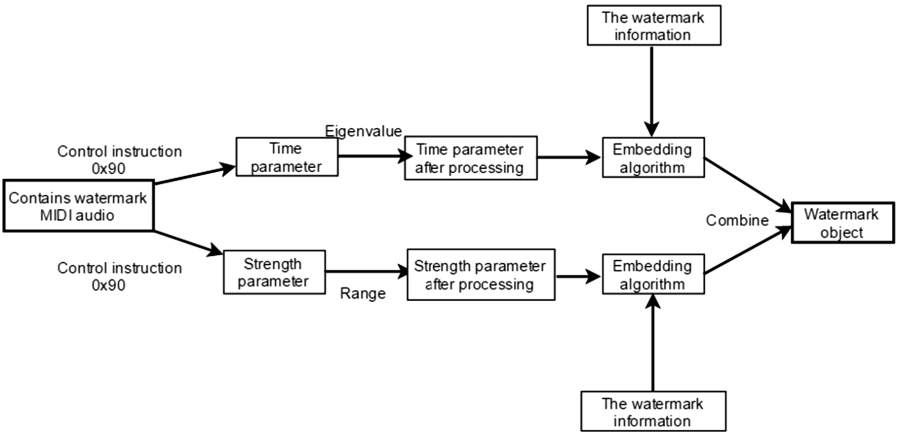


Fig. 1. MIDI basic block diagram of the audio embedding algorithm

3 Experimental Results and Analysis

At present, there are few researches on MIDI audio quality evaluation. Generally, there are two types: subjective evaluation method and objective evaluation method. Subjective evaluation method is judged by people's subjective consciousness. For the subjective evaluation of MIDI audio, ITU-T P.800 subjective evaluation method of voice quality is still used, and its essence is MOS, Mean Opinion Score.

In order to improve the shortcomings of the subjective evaluation method, people put forward some objective evaluation methods about the quality of audio, and the objective evaluation method is to make a reasonable evaluation of the audio quality by simulating the human auditory system.

3.1 Objective Evaluation Method

The mean square error is a convenient method to measure the "mean error", which can evaluate the degree of change of the data. The smaller the MSE value, the better accuracy of the prediction model is described.

The MSE method is described as follows:

$$MSE = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}(g(x,y) - g'(x,y))^2}{M \times N} \tag{1}$$

$g(x,y)$ is the original MIDI audio, $g'(x,y)$ is embedded after watermark MIDI audio, is the number of rows of MIDI audio data, is the number of columns of MIDI audio data.

PSNR is the most widely used and most common objective evaluation index of image. The PSNR method of image is based on the error between corresponding pixels, that is an image quality evaluation based on error sensitive. Combining with the characteristics of MIDI audio dynamic parameters, Yang [5] improved the PSNR method of image quality to get a PSNR method to measure the degree of the audio change to the original MIDI.

$$PSNR = -10\lg\left\{\frac{1}{127}\sum_{n=1}^{N} d(n)^2\right\} \tag{2}$$

N is the number of original MIDI audio intensity values, $d(n)$ represents the difference between the initial strength and the initial strength of the n intensity value, the sum in parentheses is the total amount of energy that the cryptography causes distortion, is the peak signal energy, the mean of each these forces is equal to the total energy at the peak of 127.

The Included Angle Cosine method is described as follows:

The angle cosine is used to measure the angle between the two groups of vectors, also known as the sum coefficient. The expression as follow:

$$\cos(x,\ y) = \frac{(x,\ y)}{||x|| \cdot ||y||} = \frac{\sum_{i=1}^{n} x_i y_i}{\left(\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2\right)^{1/2}} \tag{3}$$

In which x_i is the original MIDI audio, y_i is MIDI audio embedded in secret information. The geometric meaning of cosine is the cosine value of the angle between the two vectors in the N dimensional space composed of N elements. In general, the elements in the vector need to be dimensionless before use, and all the elements are positive. The value range of the angle cosine is [0, 1], the larger the value, the smaller the two vector angle, the closer the two vectors, the value of the 1, the two vectors are exactly the same.

3.2 Experimental Data and Analysis

Table 1. MSE method

Time parameter	MSE	Strength parameter	MSE
Change the lowest 1bit	0.0589	Change the lowest 1bit	0.0379
Change the lowest 2bit	0.2687	Change the lowest 2bit	0.2743
Change the lowest 3bit	0.7909	Change the lowest 3bit	0.8942
Change the lowest 4bit	2.0467	Change the lowest 4bit	2.9892

Table 2. PSNR method

Embed 0, 1 secret information randomly	PSNR (dB)	All embedded 1as secret information	PSNR (dB)
Change the lowest 1bit	45.0883	Change the lowest 1bit	42.07
Change the lowest 2bit	38.4536	Change the lowest 2bit	34.36
Change the lowest 3bit	33.1221	Change the lowest 3bit	28.77
Change the lowest 4bit	28.0180	Change the lowest 4bit	23.41

Table 3. Angle cosine method

Time parameter	Included angle cosine	Strength parameter	Included angle cosine
Change the lowest 1bit	0.9998	Change the lowest 1bit	1
Change the lowest 2bit	0.9988	Change the lowest 2bit	0.9999
Change the lowest 3bit	0.9961	Change the lowest 3bit	0.9996
Change the lowest 4bit	0.9890	Change the lowest 4bit	0.9986

Table 1 shows that the MSE method can significantly perceive small changes in the audio quality of MIDI, from Table 2, it can be found that the quality of MIDI audio loudness changes, as shown in Table 3, it can be seen that the Angle cosine method is sensitive to large changes in MIDI file instructions. At present, PSNR method is not mature and can only evaluate the change range of MIDI audio intensity parameters.

4 Conclusion

In this paper, in order to increase the MIDI audio embedding capacity problem, improve the original LSB algorithm that only for single parameter embedding watermarking, put forward a kind of time and strength of the two parameters and steganalysis algorithm, using the statistical properties of the parameters of a blind extraction threshold value, which can realize blind extraction, and can reduce the embedded secret information on the quality of the original audio, enhances the ability

of resisting steganographic analysis at the same time. Based on the existing image quality evaluation methods, three objective quality evaluation methods of MIDI audio are proposed. The experimental results show that the effect of the algorithm on the sound quality is almost negligible, and the ability of anti-steganography analysis is greatly improved. Compared with the traditional MIDI steganography algorithm, the hidden writing capacity of this algorithm is obviously improved, and the damage to the MIDI sound quality is very low.

Acknowledgments. 1. This research project was supported by the Education Department of Jilin Province ("Thirteen Five" Scientific Planning Project, Grant No. 249 [2016]).

2. This research project was supported by the National Natural Science Foundation of China (Grant No. 61262090).

References

1. Wu, Y.: A research on three kinds of watermarking technology. Master Thesis of Harbin Engineering University (2004)
2. Yang, B.: Basic application of MIDI music is briefly discussed. Sci. Technol. Innov. Herald. 246–246 (2010)
3. Chen, X., Liu, Y., Pan, X., Ma, J.: The Development and Application of MIDI Principle, pp. 34–38. National University of Defense Technology Press, Beijing (2008)
4. Yang, B.: MIDI audio steganography and steganalysis. Master Thesis of University of Science and Technology of China, pp. 19–25 (2010)
5. Yang, F.: Research on steganalysis in MIIDI audio. Master Thesis of University of Science and Technology of China, pp. 11–15 (2009)
6. Yang, Y.: An experimental course on information hiding and digital watermarking. National Defense Industry Press, Beijing (2012)

WSN/RFID Indoor Positioning and Tracking Based on Machine Learning: A Health Care Application

Ermal Elbasani, Hyun Lee, and Jae Sung Choi(✉)

Department of Computer Engineering, Sun Moon University,
Asan-si, ChungCheongnam-Do 31460, Korea
ermal.elbasani@gmail.com,
{mahyun91, jschoi}@sunmoon.ac.kr

Abstract. RFID and WSN technologies have revolutionized monitoring and tracking systems for healthcare domain in indoor applications as a current trend and challenge at the same time for Internet of Things. WSN, RFID and Machine Learning have been widely applied as a promoter of Smart Homes to opened novel ways for research and exploration in wellness and healthcare industry. The main focus of this study is to related location and positioning in a Smart Home system to support real time monitoring of weak and elderly in order that smart home devises to respond their needs. Based on this observation this research aims to provide a better accuracy on indoor tracking and localization by using location dataset and we have implemented Machine learning methods Naïve Bayes, Support Vector machine SMV and archived a location accuracy of 99% and 98% respectively.

Keywords: RFID · WSN · Machine learning · Indoor localization

1 Introduction

In smart indoor environment, Localization is a useful and must-have feature for knowledge discovery, especially those who need long-term monitoring, to discover patterns about behaviors and other request of the users. Most of the research related to indoor localization topics and healthcare has brought results based on devices which determined the location by sensing the signals by the so-called BSN (Body Sensor Network) wearables or RFID (Radio-Frequency Identification) methods. In order to development an intelligent health system with the main functionalities of tracking and identification requires the integration of RFID and Wireless Sensor Networks (WSN) that's provides a low-cost and scalable system.

Two main drawbacks in existing smart health solutions: first, the lack of integration between different Internet of Things (IoT) technologies; and second, the lack of implementations in real scenarios including integration with the Hospital Information System (HIS). Based on real practical project of WSN-RFID integration in health care system [1] and selecting the recent approaches of machine learning technologies [2, 3] derivate to an issue that is localization in real time. Real-time localization allows

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 446–452, 2020.
https://doi.org/10.1007/978-981-13-9341-9_77

patients or other employees to quickly find health care providers in large-scale facilities when seeking help in an emergency. By connecting to a specific location, detailed data near the patient can be provided at a specific time. In addition, elderly nursing homes can provide information on user's mobility through the RTLS (Real-time Localization System), which is an of resident's overall well-being. Our approach can be considered as an extension of similar studies [1], that have implemented KNN, ANN and Linear regression methods for achieving a high accuracy in localization. We used this study in order to be able to use in another scenario for indoor localization. In our approach, we implementing two Machine Learning methods SVM and Naïve Bayes outperforming the previous studies using an architecture as shown in Fig. 1.

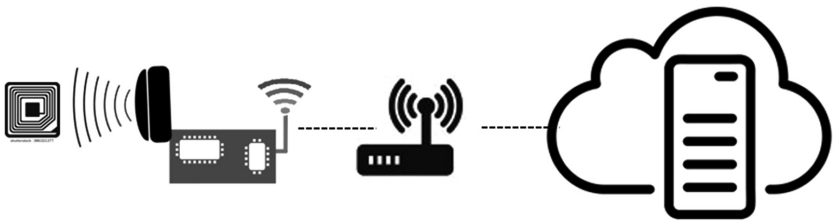


Fig. 1. Architecture of a hybrid RFID/WSN monitoring

2 Related Work

In localization and positioning systems there is widespread attention and expectations in academic and industrial research, and many technologies have recently been used and combined. In many medical applications positioning and tracking are an important feature, since knowledge of the accurate location of people and assets is necessary [8]. Based on the RSS (Received Signal Strength) of a single wireless technology, have been made several efforts to develop a positioning system, and in order to solve this problem, methods such as K-Nearest Neighbors (KNN), GPS and other triangulation techniques, Kalman Filtering and neural network are used [9]. Machine learning method implemented on similar example achieve an average of 2 m of accuracy, but the current research includes one or a few algorithms, most of which do not use many sensors available in smart mobile devices, and also the size of the data is another issue associated. In our scenario is integrated RFID and WSN methods, we use machine learning techniques to calculate location, and we tend to get better position accuracy with fast processing time.

3 Architecture Overviews

Many wireless communication technologies are widely used to develop indoor positioning systems. These technologies include RFID, Bluetooth and WSN and wireless local area network (WLAN). In order to improve accuracy for indoor localization and

positioning, we propose to integrate the machine learning methods exploiting the data form different modules indoor positioning system.

UHF (ultra-high-frequency) and RFID is highly recognized as an important technology in detecting and identifying systems, furthermore widely regarded as a potential candidate for localization in the IoT. In the positioning system based on RFID, the CoO (Cell of Origin) positioning method is mainly adopted, but with this positioning method, the accuracy is dependent on number of tags and the maximum tag-reading range. Other RIFD-based localization method highlights RSS which is applied in multilateration positioning methods. Phase of arrival positioning is another accurate method widely use in WSN but also have integrated to localize object only based RFID tags and readers [7]. The more recent developments, WSN sensors embedded with RFID tag, and the identification and tracking are process through the network. The most supportive and aforementioned techniques of WSN, are IEEE 802.15.4, ZigBee and Wi-Fi or Bluetooth also.

Fingerprinting in localization are data-driven approaches that function with the assumption that specific RF features are considered to be data points that identifies a location in a unique way. Fingerprinting algorithms can be installed on top of available infrastructures, with minimum requirements, and are a very useful option for fast indoor localization environments. The disadvantage is that when the environment changes, the training database need to be updated or sometimes even renewed in order to track the fingerprints. There are many studies involving theoretical cases of the fingerprinting printing method. Among machine learning method in this field, the most noticeable are the so-called predictive based training methods, neural network methods, K-means and also clustered algorithms [3].

Machine learning algorithms search for patterns and rules in a given data and are widely used in a variety of applications and platforms. They use the example to generate specific features and automatically learn from the data. The work of machine learning methods is generally constructed in two phases. The first phase of the learning, is training phase, as data is collected and processed to be computed from algorithm, patterns can be learned, extract features and create the classifying models. In the second phase of learning is the testing phase, new data entered is tested regarded the model created during the training phase, and the efficiency of the model is then calculated for the accurate result (Fig. 2). This training-testing of learning methods belong to so-called a supervised learning algorithm. In order to improve the indoor localization prediction function, it is recommended to integrate mechanical learning technology compared with the previous work. In this work the localization is based using two supervised machine learning methods Naïve Bayes and Support vector machine (SVM).

Naïve Bayes: In study [4] a system developed for positioning WSNs using the Bayesian algorithm including rather few number of anchor nodes. This study focuses on the improvement of gradual corrections, which is a method of predicting models from likelihoods, bringing it closer to posterior likelihood. This machine learning method have shown effective results for localization with large number of nodes.

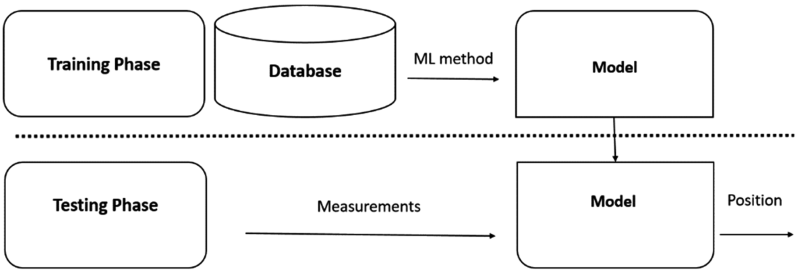


Fig. 2. ML schema for training and test dataset for the localization.

Support Vector Machine (SVM): SVM technology is widely used in WSN positioning, and specific cases [5, 6] where for nodes or sensor in not easily practically to self-position. In research [5] there is a model of mobile localization using functionalities of SVM. The proposed approach in the early stages is to use wireless frequency oscillations like RSS metric to sense node movement. To provide a new location for mobile awareness, the SVM will be executed.

4 Method

To evaluate proposed ML methods there is the UJIIndoorLoc dataset [11] that contains Wi-Fi measurements used during EvAAL competition at [10] and is made publicly available. The source dataset itself is created based of Jaume I University from WAP signal information in interaction with smartphones. In our perspective we use RFID technology to replace the phones and still keep the ID is received from the measures. Machine learning is going to use the prediction method of the trained dataset and test dataset in order to give how much does the test dataset approach to the trained one. Other procedure is arranging dataset from source and splitting dataset for training and test dataset. The first step was to remove columns and rows where WAP signal strength has only one value, meaning that the WAP was not detected. Also, the columns were converted into appropriate data types. Original dataset used WAP signal range from −104 to 0, and value 100 when WAP was not detected. I have changed the values so that 0 represents the value when WAP was not detected and the highest signal is 104. We also reduced the dataset, using only some part of it, to make it similar to smaller living environment. After the regulations, we used Kernel SVM since for linear SVM does not effort a good performance.

Table 1. Performance of machine learning methods

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)
SVM	98	97	99
Naïve Bayes	99	98	99

Table 2. Evaluation of machine learning performance

Measure	Formula	Description
Accuracy	$\frac{TP + TN}{TP + TN + FP + TN}$	Overall effectiveness of a method
Sensitivity	$\frac{TP}{TP + FN}$	Effectiveness of a method to identify positive labels
Specificity	$\frac{TN}{FP + TN}$	How affectively a method identifies negative labels

5 Experimental Results

The performance of proposed calculation for localization method is evaluated through simulation. We have used R Studio environment to analyze the datasets. Speaking for the impact of the Data quality for the location method for sure we have to explain that with the arranged dataset outperform on around 10% in accuracy for both algorithms (Table 1). When we want to compare the performance of the methods, we may explain that compare with KNN and Linear Regression method on [12] we have achieved a better result. In Fig. 3(a) is shown the performance of the classification is computed based on the research's cases based on binary classification. the performance s by evaluating the TP, TN, FP, FN classes as explained [13] and formulated in Table 2. Based only on the computer selection features we may say that the result gives a better result and in organized data only for localization Kernel SVM and Naïve Bayes. For the execution time Fig. 3(b) for the same conditions Naïve Bayes shows to offer less complexity and more likely to be used in WSN. We want to mention that our study does not include real life calculation and measures, so regarded to energy consuming in a network also is not calculated. Results of the same calculated may be interpreted in detecting position of a user in less than 1 m, and even to cm mistake.

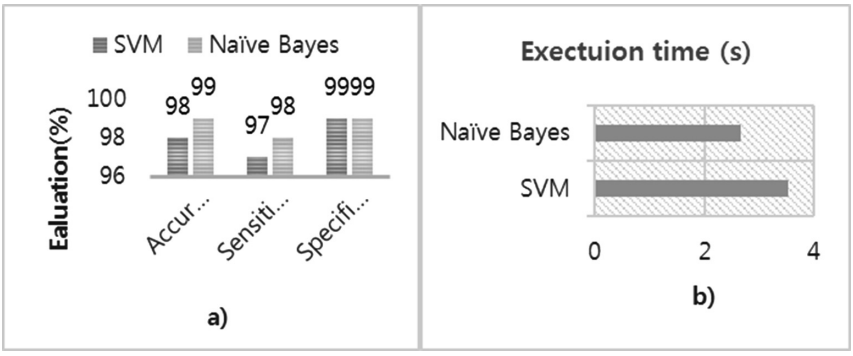


Fig. 3. (a) Performance of machine learning methods, (b) Maximum execution time

6 Conclusion

The main motivation of this study is to show that how much the machine learning methods can contribute in many of the technologies. Location awareness is part of many researches to address many issues especially in health care and wellness, when is needed to monitor people in physical handicap. Not just monitor but as a future project of Smart Homes to integrate many devices in order to serve the person in need. Features to be classified are selected from experienced technicians and in laboratory experiments. So, in a real scenario the computer selected features this kind of performance will be not measured again or should have the cost of other calculations in the network that can affect the energy and also communication.

In order to achieve a better positioning accuracy or improve performance of the machine learning method when using more training data or having a proportional anchors node in the network. The future research topic is to deal with the scalability of data from network. Considering for research purpose, in our project, the number of sensor nodes is sufficient for training stage. Furthermore, it will be interesting to study how to achieve results also in beacon mobile nodes and learning how will affect the positioning and tracking in indoor environment

Acknowledgments. This work is result of studies on the "Leaders in INdustry-university Cooperation" Project, which is supported by the Korean Ministry of Education, and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2018R1C1B5045953)

References

1. Adame, T., Bel, A., Carreras, A., Seguí, J.M., Oliver, M., Pous, R.: CUIDATS: an RFID–WSN hybrid monitoring system for smart health care environments. Future Gener. Comput. Syst. **78**, 602–615 (2016)
2. Ahmadi, H., Bouallegue, R.: Exploiting machine learning strategies and RSSI for localization in wireless sensor networks: a survey. In: International Wireless Communications and Mobile Computing Conference IEEE (2017)
3. Alsheikh, M.A., Lin, S., Niyato, D., Tan, H.P.: Machine learning in wireless sensor networks: algorithms, strategies, and applications. IEEE Commun. Surv. Tutor. **16**(4), 1996–2018 (2014). Fourthquarter
4. Morelande, M., Moran, B., Brazil, M.: Bayesian node localization in wireless sensor networks. In: Acoustics, Speech and Signal Processing, pp. 2545–2548. IEEE (2008)
5. Yang, B., Yang, J., Xu, J., Yang, D.: Area localization algorithm for mobile nodes in wireless sensor networks based on support vector machines. In: Mobile Ad-Hoc and Sensor Networks, pp. 561–571. Springer (2007)
6. Kim, W., Park, J., Kim, H.: Target localization using ensemble support vector regression in wireless sensor networks. In: Wireless Communications and Networking Conference, pp. 1–5 (2012)
7. Povalac, A., Sandebesta, J.: Phase of arrival ranging method for UHF RFID tags using instantaneous frequency measurement. In: Conference Proceedings ICECom, pp. 1–4. IEEE Press, Piscataway (2010)

8. Hightower, J., Borriello, G.: A survey and taxonomy of location systems for ubiquitous computing. Technical report UW-CSE 01-08-03 (2001)
9. Gante, A.D., Siller, M.: A survey of hybrid schemes for location estimation in wireless sensor networks. Procedia Technol. **7**, 377–383 (2013)
10. Moreira, A., Nicolau, M.J., Meneses, F., Costa, A.: Wi-Fi fingerprinting in the real world - RTLS@UM at the EvAAL competition. In: (IPIN), pp. 1–10 (2015)
11. Sospedra, J.T., Montoliu, R., Usó, A.M., Avariento, J.P., Arnau, T.J., Bordonau, M.B., Huerta, J.: UJIIndoorLoc: a new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems. In: International Conference (IPIN), pp 261–270 (2014)
12. Ugave, V.A., Pasricha, S.: Smart Indoor Localization Using Machine Learning Techniques. Colorado State University, Fort Collins (2014)
13. Sokolova, M., Lapalme, G.: A systematic analysis of performance measures for classification tasks. Inf. Process. Manag. **45**, 427–437 (2009)

Parking Occupancy Detection: A Lightweight Deep Neural Network Approach

Chin-Kit Ng(✉), Soon-Nyean Cheong, and Yee-Loo Foo

Faculty of Engineering, Multimedia University,
63000 Cyberjaya, Selangor, Malaysia
chinkitng@yahoo.com, {sncheong,ylfoo}@mmu.edu.my

Abstract. Inaccessibility of real-time parking occupancy information may cause inefficiency in parking management. This paper proposed a novel lightweight deep neural network approach to realize outdoor parking occupancy detection system to support more efficient parking management. A lightweight MobileNet binary classifier is used to accurately identify the occupancy status of parking space image patches that are extracted from live parking lot camera feeds. A performance comparison between different network configurations of MobileNet has been done to investigate their speed-accuracy trade-off when running on embedded device. The prototype was deployed at an outdoor campus parking to evaluate effectiveness of the proposed system. The prototype can detect 22 parking spaces within 2.4 s when running on an ASUS Tinker Board and achieve a detection accuracy of 99%.

Keywords: Lightweight deep neural network · Parking occupancy detection

1 Introduction

Poor parking management is often due to the unavailability of reliable parking occupancy information which is crucial to understand the utilization patterns of parking assets. As such, growing interest is observed among researchers to actively explore smart parking solutions to better manage parking resources. Recent researches exhibit rising trend in applying computer vision technology for parking occupancy detection rather than using parking detection sensors because vision-based systems can provide additional security surveillance while having lower implementation costs.

Convolutional Neural Network (CNN) is the image recognition algorithm with state-of-the-art performance on the ImageNet challenge [1]. Many deeper and more complicated networks have been built to improve the accuracy performance at the cost of increasing network size and computational complexity. However, real-time parking occupancy detection usually requires image classification task to be performed on resource limited embedded devices. Fortunately, a lightweight deep neural network named MobileNet [2] can potentially meet the requirements of low latency application. It consumes much less computational cost to run with negligible accuracy trade-off for simple classification tasks. This research work proposed the novel use of MobileNet in detecting live occupancy status of outdoor parking spaces. It has been implemented at an outdoor parking of Multimedia University, Malaysia. This paper contributes to the

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 453–458, 2020.
https://doi.org/10.1007/978-981-13-9341-9_78

better understanding on the speed-accuracy trade-off of different MobileNet configurations and validate their applicability for low latency parking occupancy detection on single board computer.

2 Related Work

Background subtraction is one typical approach to identify parked vehicles in the foreground from parking lot images. In the work by Liu et al. [3], presence of vehicle is determined by analyzing the edge density, closed contour density and foreground/background pixel ratio of each parking space. The detection mechanism presented in [4] determines occupied spaces by converting parking lot image into binary form followed by background elimination to check for the existence of vehicle. However, both works in [3] and [4] do not provide measures to deal with different weather conditions which could deteriorate the detection performance.

A recent work by Nieto et al. [5] used a trained vehicle detector based on the Faster R-CNN [6] to locate vehicles in video frame. A spot mapping algorithm then automatically maps each detected vehicle to the corresponding spaces in the parking grid. This deep learning-based method offers superior robustness against variable background caused by lighting and weather changes. Nevertheless, such approach requires GPU acceleration to ensure reasonable detection speed. This limitation makes the solution infeasible to operate on embedded devices that are preferable in constructing a decentralized system where each parking area is installed with a dedicated detection unit to prevent single point of failure.

In the work by Amato et al. [7], CNN classifier is employed for parking occupancy detection on embedded devices. The proposed technique demonstrated robustness towards weather changes in its performance benchmark on PKLot [8]. Another work combined trained CNN with background subtraction to improve parking space status classification performance on radar image [9]. Both the CNN-based approaches require network customization of large-sized CNN model to obtain a sufficiently small network that can run on computationally limited platforms.

3 System Design

An overview of the proposed parking occupancy detection system is depicted in Fig. 1. An outdoor IP camera is locally connected to a base station to collectively function as the detection unit. Prior to system deployment, CNN training procedures are carried out to obtain a trained image classification model. Firstly, parking lot images are collected for a couple of weeks under changing weather conditions. Image patches of individual parking space are extracted from the collected parking lot images based on an annotations file specifying the coordinates of each parking space. The extracted image patches are manually divided into two classes namely occupied and empty to form the training image dataset which is then used for the CNN training. MobileNet is chosen for the training process because of its lightweight network structure that is suitable for

real-time application [2]. A trained image classification model containing optimally-tuned weight values is obtained as the training output. It will be transferred to the base station to be used for parking occupancy detection.

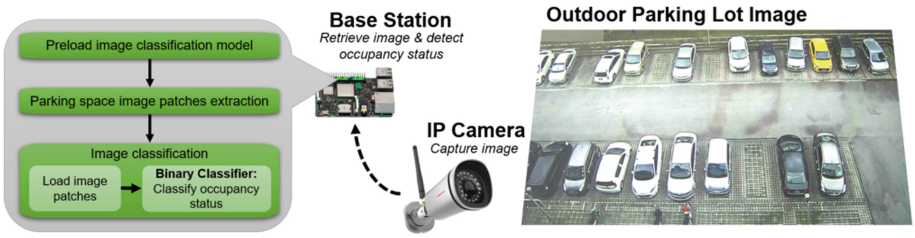


Fig. 1. Architecture of parking occupancy detection system

During real-time operation, the base station continuously retrieves live parking lot image from the IP camera to enable real-time parking occupancy detection. For each detection cycle, parking space image patches are cropped out from the retrieved parking lot image. Subsequently, image classification is performed by loading the image patches and inputting them to the binary classifier which uses the preloaded image classification model to determine the occupancy status as empty or occupied.

4 System Implementation

Parking lot images (covering 22 parking spaces) are collected by taking image snapshot using a Foscam FI9800P outdoor IP camera mounted on third floor of the Faculty of Engineering building at Multimedia University. The collection of images is carried out for two weeks, starting from 8 am to 7 pm on each day with ten minutes interval. A graphical image annotation tool known as LabelImg is used to annotate and define the regions of individual parking space. Based on the annotation output file, the 22 parking space image patches are extracted out from each parking lot image using the OpenCV software library and categorized into two classes (i.e. empty and occupied) according to their respective occupancy status. Overall, around 7k image samples are yielded for each class to form the training image dataset.

Table 1. Specification of MobileNet configurations

Models	WM	Input resolution	Multi-adds computation	Parameters
Baseline MobileNet	1.00	224	568 million	3.2 million
Minimal MobileNet	0.25	128	15 million	0.2 million

The training operation of MobileNet is performed using the Tensorflow machine learning framework. To examine the speed-accuracy trade-off of different MobileNet

configurations, two hyper-parameters namely width multiplier (WM) and resolution multiplier (RM) are set to their respective maximal and minimal values such that WM, RM $\in$ (0, 1], resulting to two extreme models as shown in Table 1. The function of WM is to uniformly prune each network layer by reducing the number of input and output channels. RM is applied implicitly to a network to lower the input resolution by setting image resolution to values range from 224 down to 128. With WM set to 0.25 and input resolution of 128, the resulting minimal MobileNet has its computational cost reduced by nearly 38 times and number of parameters lessen by 16 times as compared to the baseline MobileNet. Both the MobileNet models are trained on the prepared training image dataset for 4000 iterations with initial learning rate of 0.01 and RMSprop solver type chosen. The training process completed with train accuracy and validation accuracy saturated at 100%. A trained computational graph containing the optimized trainable variables is generated as the training output which represent the image classification model.

ASUS Tinker Board is used as the base station in the proposed system to perform parking occupancy detection. It is a credit card-sized single board computer embedded with a 1.8 GHz quad-core processor and 2 GB LPDDR3 RAM. It runs on a Debian-based operating system called TinkerOS and is installed with ARM compatible Tensorflow module. The binary classifier is implemented as a Python based application with Tensorflow module imported. The detection results which specify the occupancy status of each parking space are stored in JSON format.

5 Results and Discussions

To assess detection accuracy of the proposed system, an isolated testing image dataset is prepared by collecting parking lot images for another two working days at one-minute interval. Likewise, these images exhibit interferences due to varying light intensity and shadow patterns at distinct time of a day. Parking space image patches are extracted from the collected images and segregated into empty and occupied test samples to establish the ground truth. Quantitative analysis is carried out using evaluation metrics namely False Negative Rate (FNR), False Positive Rate (FPR) and Average Detection Accuracy (ADA). A false negative (FN) means empty space is detected as occupied whereas a false positive (FP) means occupied space is detected as empty. True negative (TN) and true positive (TP) mean correct detection of occupied space and empty space respectively. The respective formula for each evaluation metric is stated as follows:

$$FPR = \frac{FP}{FP + TN} \tag{1}$$

$$FNR = \frac{FN}{FN + TP} \tag{2}$$

$$ADA = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

According to Table 2, the two MobileNet models specified in Table 1 demonstrated excellent performance for both Day1 and Day2 testing datasets, achieving 99% detection accuracy. This finding suggests that the minimal MobileNet can satisfy the accuracy requirement for outdoor parking occupancy detection despite having less parameters. It is noticed that the overall FNR for both models is significantly higher than their overall FPR (i.e. greater tendency to misclassify an empty space as being occupied). Such discrepancy is primarily due to those empty test samples that contain portion of vehicle that is leaving or entering a parking space. This error can be mitigated by applying temporal hysteresis decision scheme to allow a change in occupancy status only when the detection result is consistent for at least a few consecutive cycles. Nonetheless, such technique would cause delay of detection results in the expense of providing greater robustness.

Table 2. Accuracy of parking occupancy detection with MobileNet

			Testing dataset		
			Day 1	Day 2	Overall
No. of test samples		Empty	2791	3179	5970
		Occupied	11729	11341	23070
		Total	14520	14520	29040
Evaluation metrics	Baseline MobileNet	FPR	0.0010	0.0008	*0.0009*
		FNR	0.0201	0.0132	*0.0164*
		ADA	0.9953	0.9965	**0.9959**
	Minimal MobileNet	FPR	0.0020	0.0005	*0.0013*
		FNR	0.0355	0.0129	*0.0235*
		ADA	0.9915	0.9968	**0.9941**

Table 3. Speed of parking occupancy detection with MobileNet

Time to classify one parking space image patch (millisecond)		Speed-up ratio
Baseline MobileNet	Minimal MobileNet	
434	106	4.094

As shown in Table 3, the trained binary classifier based on minimal MobileNet can classify individual image patch in approximately one tenth of a second (i.e. around quadruple the speed of baseline MobileNet) when running on the ASUS Tinker Board. In other word, the proposed system can reflect changes of occupancy status for 22 parking spaces in less than 2.4 s.

To sum up, the experimental results evidence that the minimal MobileNet is the optimal choice for low latency parking occupancy detection on embedded device because it can provide a much higher detection speed meanwhile maintaining the same accuracy level as the baseline MobileNet. The proposed system can offer accurate parking occupancy detection which is robust against different weather conditions when compared to those background subtraction approaches [3, 4]. Such implementation

represents a more economical solution than the vehicle detector-based approach in [5] that requires GPU hardware acceleration. Besides, minimum configuration is required to prepare the MobileNet for training without the needs of manual network customization as in the previous CNN-based approaches [7, 9].

6 Conclusion

The paper presents implementation and deployment of a lightweight MobileNet vision-based outdoor parking occupancy detection system at an outdoor parking lot in Multimedia University. Compared to the baseline MobileNet, the binary classifier generated from the minimal MobileNet can enable low latency outdoor parking occupancy detection with negligible accuracy trade-off when operated on embedded device. In essence, the proposed system represents a cost-effective yet reliable smart parking solution that can be easily adopted to uplift existing parking facilities. In future research, the training dataset can be extended to include night time image samples to enable 24-h parking occupancy detection.

Acknowledgments. The authors thankfully acknowledge the financial supports provided by the Telekom Malaysia Research and Development Grant (No. RDTC170946) to successfully implement the prototype at Multimedia University (Cyberjaya, Malaysia).

References

1. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L.: ImageNet large scale visual recognition challenge. Int. J. Comput. Vis. **115**(3), 211–252 (2015)
2. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Adam, H.: MobileNets: efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861 (2017)
3. Liu, J., Mohandes, M., Deriche, M.: A multi-classifier image based vacant parking detection system. In: 2013 IEEE 20th International Conference on Electronics, Circuits, and Systems (ICECS), pp. 933–936, Abu Dhabi (2013)
4. Al-Kharusi, H., Al-Bahadly, I.: Intelligent parking management system based on image processing. World J. Eng. Technol. **2**(2), 55–67 (2014)
5. Nieto, R.M., García-Martín, Á., Hauptmann, A.G., Martínez, J.M.: Automatic vacant parking places management system using multicamera vehicle detection. IEEE Trans. Intell. Transp. Syst. **20**, 1069–1080 (2018)
6. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real time object detection with region proposal networks. In: Proceedings CVPR, pp. 91–99 (2015)
7. Amato, G., Carrara, F., Falchi, F., Gennaro, C., Vairo, C.: Car parking occupancy detection using smart camera networks and deep learning. In: 2016 IEEE Symposium on Computers and Communication (ISCC). IEEE (2016)
8. de Almeida, P.R., Oliveira, L.S., Britto, A.S., Silva, E.J., Koerich, A.L.: PKLot–a robust dataset for parking lot classification. Expert Syst. Appl. **42**(11), 4937–4949 (2015)
9. Martinez, J., Zoeke, D., Vossiek, M.: A convolutional neural network approach to parking monitoring in urban radar sensing. In: 2017 European Radar Conference (EURAD), pp. 86–89 (2017)

Detecting Driver Drowsiness Based Fusion Multi-sensors Method

Svetlana Kim[1], Hyunho Park[2], Yong-Tae Lee[2], and YongIk Yoon[1(✉)]

[1] Department of IT Engineering, Sookmyung Women's University, Yonsangu Chongpa-dong 2ga, Seoul 04310, Korea
{xatyna, yiyoon}@sookmyung.ac.kr
[2] Smart Media Research Group, Electronics and Telecommunications Research Institute, 218 Gajeong-ro, Yuseong-gu, Daejeon 34129, Korea
{hyunhopark, ytlee}@etri.re.kr

Abstract. In recent years, driver's drowsiness is one of the main causes of traffic accidents, which can result in severe physical injury and serious economic loss. Fatigue of the driver is an important factor in road accidents, and fatigue detection has a significant influence on traffic safety. This article describes a drowsiness detection approach based on the combination of various multi-sensors. The present study proposed a method to detect the driver's drowsiness that combines features of electrocardiography (ECG) and environmental factors, such as vehicle temperature and humidity, to improve detection performance. The activity of the autonomic nervous system which can be measured in heart rate variability (HRV) signals obtained from surface ECG, indicates changes during stress, extreme fatigue, and episodes of drowsiness. The combination of the multi-sensors feature of drowsiness is significant factors in determining the driver's fatigue state and can use this information to transportation drowsy driving control center if necessary.

Keywords: Sensors data fusion · Driver drowsiness detection · Biosensors · Environmental sensors

1 Introduction

Driver's fatigue detection is an important research topic in the management of driving safety systems because driver's drowsiness is a major cause of many automobile accidents around the world [1]. During the past four years (from 2013y to 2016y), according to available statistical data in Korea National Police Board has experienced 10,072 cases of drowsy driving [2]. These crashes resulted in approximately 57 deaths and 20,055 people are injuries. The German Council for Road Safety (DVR) claims that one in four highway traffic accidents is the result of the driver's immediate drowsiness [3]. These statistics show that drowsiness of the driver is one of the main causes of traffic accidents.

In recent years, a number of methods have been proposed, including subjective and objective detection methods [4–6]. However, the recognition speed based on the

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 459–464, 2020.
https://doi.org/10.1007/978-981-13-9341-9_79

method of subjective detection is highly dependent on the judgment of the driver himself or the actions of other drivers, and it is difficult to determine the driver's fatigue in real time. In drowsy driving, the driver loses control of the vehicle and often crashes with other vehicles or stationary objects [7]. To prevent these fatal accidents, that should monitor the state of drowsiness of the driver. The following four measures are widely used to monitor drowsiness.

Vehicle-based measurement based on the number of metrics are categorized into vehicle driving parameters using electronic sensors including steering wheel handling, lane tracking, and pressure on the accelerator pedal. Changes that are continuously monitored and exceed the specified thresholds indicate a significant increase in the probability that the driver is drowsy [8, 9]. *Environment measures*, such as weather, temperature in-vehicle, humidity, etc., are one of the causes drowsiness driving. If the weather is very hot or cold, the driver will use the air conditioner and this can cause drowsiness due to lack of oxygen. *Behavior measures* have detected driver behavior characteristics, such as eye closure/blinking, facial expressions, head pose, etc., is monitored through a camera using video imaging techniques [10, 11]. *Physiological measures* is a correlation between physiological signals of the driver physiological parameters, such as electrooculogram (EoG), electrocardiogram (ECG), electromyogram (EMG) and electroencephalogram (EEG) [12, 13]. The most reliable way of monitoring is the use of physiological signals such as brain waves, pulse rate, heart rate, and respiration rate signal. In this paper has investigated and proposed the ECG based classifier for driver drowsiness detection by analyzing ECG frequency band, Heart Rate (HR) and Heart Rate Variability (HRV).

2 Materials and Methods

2.1 Characteristics of ECG

The ECG is a graphical representation electrical signal activity generated by the human heart are collected through the EEG electrode. The ECG signals have a four frequency band: Delta(0.5 - 4), Theta(4-8), Alpha(8 – 12.5), Beta(12.5 – 28). It can be noted that the bands of Theta Frequency correspond to drowsiness and are of great interest. When the person is sleeping, the frequency range of the theta band shows a significant increase in the power spectrum. These variations are recorded for analysis, and once the threshold crosses, the driver is labeled drowsy.

The variations in Heart Rate (HR) are visibly displayed when switching from a state of readiness to drowsiness. Each heart cycle generates ECG waves labelled P, Q, R, S, and T to represent different phases of the heartbeat. Heart rate is the number of heartbeats per unit of time, expressed as beats per minute (bpm), which is the frequency of heart rate change (HRV). The HRV analysis provides numerous frequency spectra, use two frequency bands with physiological significance: Low-frequency (LF) band (0.04–0.15 Hz) and High- frequency (HF) band (0.14–0.5 Hz). The power of the LF band depends on the sympathetic and parasympathetic activity. The high frequency (HF) band is considered as a parasympathetic source in the classic HRV analysis when the respiratory rate is considered to be in the range of 0.15 to 0.4 Hz. The ratio of LF to

HF in the ECG gradually decreases when the driver progresses from wakefulness to a drowsy state. [14, 15].

Measurement of raw physiological signals is always subject to noise and artifacts. Therefore, various pre-processing techniques such as a low-pass filter and a digital differentiator are used to remove noise. In general, effective digital filtering technology will eliminate the undesirable artifacts in an optimal way. Extract multiple statistical functions from processed signals using various characterization techniques, including Fast Fourier Transform (FFT) and Discrete Wavelet Transform (DWT).

In this paper, by analyzing the effect of multiple entropy fusion, we propose an ECG based model that can effectively detect the driver's fatigue condition. ECG-based systems may detect fatigue in relevant areas and may play an additional role in existing methods. Figure 1 shows the operation of an ECG-based fatigue detection system including data collection and preprocessing, data modeling and processing of feature extraction technical innovation, and model output analysis results.

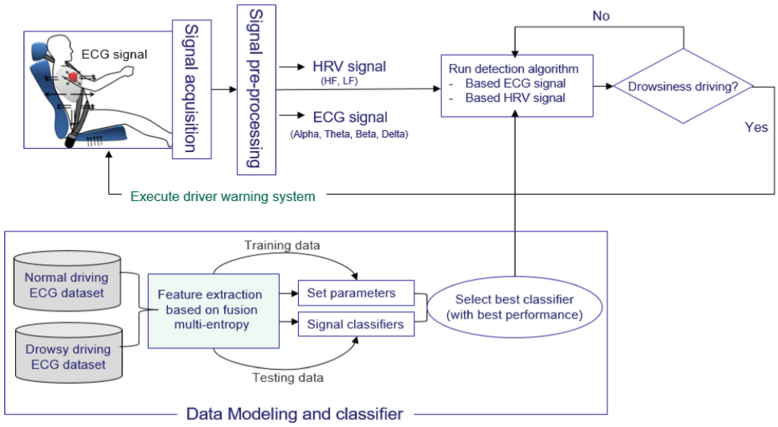


Fig. 1. A Flowchart shows the operation process of ECG-based drowsy detection.

2.2 Feature Extraction and Classification

Analysis of the time domain of ECG data does not provide detailed information on the relationship between the activation of physical conditions (i.e., the frequency band that is activated when drowsiness) and the physiological state of a person. Thus, the frequency domain of ECG data is analyzed, which includes absolute power and relative power in all channels and frequency bands of the ECG. The ECG frequency band was divided into frequency bands: Delta, Theta, Alpha, and Beta. For all frequency bands, the absolute power was calculated across each electrode:

$$\text{Delta}(F_{\delta-abs}), Theta(F_{\theta-abs}), Alpha(F_{\alpha-abs}), Beta\left(F_{\beta-abs}\right).$$

Analysis of the HRV signal obtained from an untreated ECG signal was also performed and divided into three frequency bands; low frequency (LF:0.04–0.15 Hz)

and high frequency (HF: 0.15–0.4 Hz) and a very low frequency (VLF: below 0.04 Hz). The ratio of signal LF to signal HF was also evaluated, which ensured an even distribution of body weight. Studies of normalized signals in HRV are important because they help to understand behavior that balances sympathetic and parasympathetic activity. The extracted features are then classified using Support Vector Machines (SVM), *K*-nearest neighbor (*K*NN), and Random forest (RF).

2.3 Drowsiness Detection

The flowchart in Fig. 2 shows a structure consists of two operations for determining the onset of drowsiness. The operation 'A' then calculates the drowsiness index based on frequency bands of ECG, the operation 'B' measures HRV and analyzes HF and LF/HF. The next step compares the frequency bands of the ECG and HRV to calculate the possibility of drowsiness. If this function determines that drowsiness is unavoidable, it will interfere with the function of the vehicle control system to warn the driver of the danger level or avoid traffic accidents.

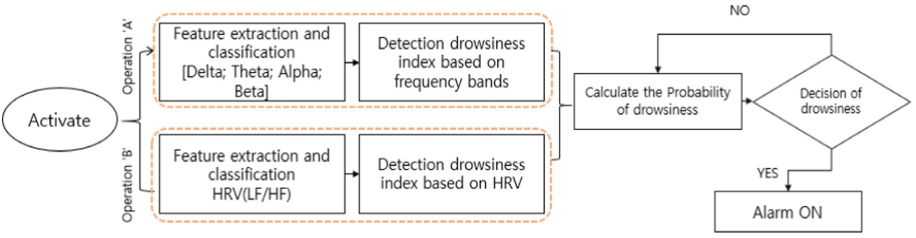


Fig. 2. A Flowchart shows the function of drowsiness detection

3 Experimental Results

3.1 ECG Datasets and ECG Data Collection

The PhysioBank Databases is a large and growing archive of physiological data. To construct the driver drowsiness classifier and detection we use a 'Real-world ECG driving datasets from Stress Recognition in Automobile Drivers Database' set. There are 18 records capturing from actual driving on highways and city. For ECG data collection we using BMS-AE-DK (biomedical system development kit) biosensors in (http://www.hybus.net/goods/view.asp?idx=89&category=25) are repeatedly for data collection.

3.2 Physiological Measures

The results of this study show that the most important parameter in terms of drowsiness classification is the HRV analysis parameter: Low-frequency and high-frequency bands - Tachogram power ratios and total power distribution in the LF and HF bands. The analysis of ECG data showed in Fig. 3, that the functions of HRV differ significantly between awake and drowsy states.

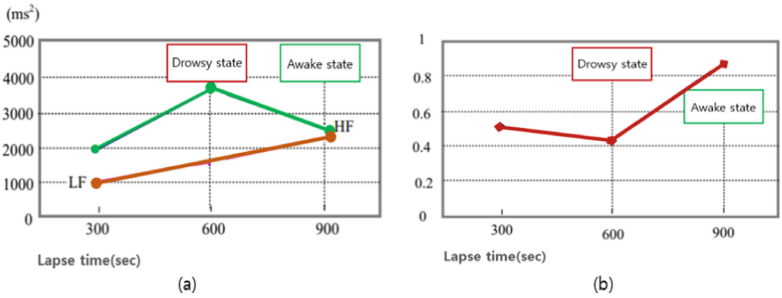


Fig. 3. (a) LF, HF heart rate variability (b) LF/HF heart rate variability

One of the goals of this study is to be able used in various physiological signals to increase the effectiveness of the driver's drowsiness detection system. First, we examined the effect of the ECG frequency band and the HRF signal separately. Then used the SVM to see the effect of using two types of signals for the efficiency of the system to detect drowsiness by combining important functions. The performance indicators confirmed that the classification accuracy achieved with the ECG frequency was 68%, the accuracy by the HRF function was 73.26%, and the combination of the two types achieved was 74.40%. This conclusion leads to the fact that combining the set of various physiological signals greatly improves the performance of the system.

4 Conclusions

In this article, a classifier based on an electrocardiogram was proposed to determine the drowsiness of the driver. To demonstrate the effectiveness and reliability of the proposed method, four common classifiers were used to learn and test the data. To detect the drowsiness of the driver using physiological data, two operations have been developed. The heart rhythm (HR) algorithm is based on a change in the heart rate and has a rapid response. Compare approaches for different states of awarening and drowsiness from real actual ECG data available in the PhysioBank archive. The result shows that the heart rate and the LF/HF ratio drop when a driver enters the drowsy mode. However, in the future needed to solve some issues. The experiment needs to test with actual real-world driving ECG data. Future research should focus on determining the driver's fatigue in real time by comparing efficiency and fusion of other frequency bands with other measurements (i.e., based on vehicle and environmental factors).

Acknowledgments. This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funder by the Korea government (MSIT) (2017-0-00336, Platform Development of Multi-log based Multi-Modal Data Convergence Analysis and Situational Response.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2018R1D1A1B07047112).

References

1. Seugnet, L., Boero, J., Gottschalk, L., Duntley, S.P., Snaw, P.J.: Identification of a biomarker for sleep drive in flies and human. Proc. Natl. Acad. Sci. **103**(52), 19913–19918 (2006)
2. http://www.police.go.kr/portal/bbs/list.do?bbsId=B0000011&menuNo=200488
3. Husar, P.: Eyetracker Warns against Momentary Driver Drowsiness. http://www.fraunhofer.de/en/press/research-news/2010/10/eye-tracker-driver-drowsiness.html. Accessed 27 July 2012
4. Kosmadopoulos, A., Sargent, C., Zhou, X., Darwent, D., Matthews, R.W., Dawson, D., et al.: The efficacy of objective and subjective predictors of driving performance during sleep restriction and circadian misalignment. Accid. Anal. Prev. **99**, 445–451 (2015)
5. Schmidt, E., Decke, R., Rasshofer, R.: Correlation between subjective driver state measures and psychophysiological and vehicular data in simulated driving. In: Intelligent Vehicles Symposium (IV), pp. 1380–1385. IEEE (2016)
6. Chen, C., Li, K., Wu, Q., Wang, H., Qian, Z., Sudlow, G.: EEG-based detection and evaluation of fatigue caused by watching 3DTV. Displays **34**(2), 81–88 (2013)
7. Wierwille, W.W., Ellsworth, L.A., Wreggit, S.S., Fairbanks, R.J., Kirn, C.L.: Research on vehicle-based driver status/performance monitoring development, validation, and refinement of algorithms for detection of driver drowsiness. Technical report, DOT HS 808 247, Office of Crash Avoidance Research National Highway Traffic Safety (1994)
8. Forsman, P.M., Vila, B.J., Short, R.A., Mott, C.G., van Dongen, H.P.A.: Efficient driver drowsiness detection at moderate levels of drowsiness. Accid. Anal. Prevent. **50**, 341–350 (2012)
9. Sahayadhas, A., Sundaraj, K., Murugappan, M.: Detecting driver drowsiness based on sensors: a review. Sensors **12**(12), 16937–16953. https://doi.org/10.3390/s121216937, PMID: 23223151 (2012)
10. Zhang, Z., Zhang, J.: A new real-time eye tracking based on nonlinear unscented Kalman filter for monitoring driver fatigue. J. Control Theor. Appl. **2010**(8), 181–188 (2010)
11. Jo, J., Lee, S.J., Park, K.R., Kim, I.-J., Kim, J.: Detecting driver drowsiness using feature-level fusion and user specific classification. Expert Syst. Appl. **41**(4), 1139–1152 (2014)
12. Jianfeng, H., Zhendong, M., Ping, W.: Multi-feature authentication system based on event evoked electroencephalogram. J. Med. Imaging Health Inform. **5**(4), 862–870 (2015)
13. Wang, H.: Detection and alleviation of driving fatigue based on EMG and EMS/EEG using wearable sensor. In: 2015 Proceedings of the 5th EAI International Conference on Wireless Mobile Comm. and Healthcare. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), pp. 155–157 (2015)
14. Task Force of ESC and NASPE: Heart rate variability: standards of measurement, physiological interpretation, and clinical use. Circulation **93**(5), 1043–1065 (1996)
15. Patel, M., Lal, S.K.L., Kavanagh, D., Rossiter, P.: Applying neural network analysis on heart rate variability data to assess driver fatigue. Exp. Syst. Appl. **38**, 7235–7242 (2011)

Overview of Data Deduplication Technology in a Cloud Storage Environment

Won-Bin Kim and Im-Yeong Lee(✉)

Department of Computer Science Engineering, Soonchunhyang University, Asan-si 31538, Republic of Korea
{wbkim29, imylee}@sch.ac.kr

Abstract. Data deduplication technology involves improving the data storage efficiency while storing and managing large amounts of data. Data deduplication reduces storage requirements by determining whether replicated data is being added to the storage, and thereby omitting the data upload if the same data already exists. Data deduplication requires data confidentiality and integrity when being applied in a cloud storage environment, and therefore various security measures such as encryption are needed. However, common encryption technologies generally cannot be applied simultaneously with data deduplication owing to the inherent inability to transform the source data. Various studies have been conducted for addressing the said issues. This paper discusses the basic environment of data deduplication technology. It also analyzes and compares multiple proposed techniques for addressing security threats.

Keywords: Date deduplication · Security · Encryption · Cloud storage

1 Introduction

Cloud storage is characterized by the provision of remote storage over a network, which allows multiple users to access and use it simultaneously. Therefore, since various data are added, it is necessary to always reserve available storage space. Data deduplication technology has been devised to prevent the storage of the same data repeatedly to save storage space.

Data deduplication technology is a technology designed to efficiently store data in a storage space. This technique uses a method to prevent the repeated storage of the same data to improve the storage efficiency of the data. However, there are various requirements for each computing environment, and different technologies must be applied to meet them. Typically, encryption technology can be mentioned. A cloud storage environment is an environment in which storage is available remotely over a network. In general, a remote server is treated as a semi-trusted environment, and accordingly, it is necessary to take appropriate measures to account for data leakage by insiders or external parties [1]. Therefore, the data stored in cloud storage must be encrypted, and at the same time, contain no information that can recover the source data. In this way, data deduplication technology is important to apply appropriate

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 465–470, 2020.
https://doi.org/10.1007/978-981-13-9341-9_80

technology considering various environments. In Sect. 2, we discuss the types of data deduplication techniques, Sect. 3 explains the types of data deduplication techniques and security issues, and Sect. 4 discusses trends in deduplication systems using these techniques.

2 Data Deduplication

This chapter looks at the various forms and application technologies that can perform deduplication on cloud storage.

2.1 Classification of Deduplication Site

Data deduplication on cloud storage is classified into client-side deduplication, appliance deduplication, and server-side deduplication, depending on the deduplication site of the data.

The way to receive the source of the file and handle deduplication in cloud storage is called server-side deduplication. This method is characterized in that data deduplication processing can be performed without imposing an additional burden on the user. However, since all the source data that is not deduplicated is transmitted, if the communication environment is poor, the user may have a heavy burden to upload the data.

A client-side deduplication (CSD) technique is a method designed for this purpose. The client side deduplication method is a method in which a list of data stored in the cloud storage is checked on the user's device and only the data that is not stored is transmitted. Therefore, the higher the overlap ratio between the data to be uploaded and the data stored in the server, the less data is uploaded, so it is easy to remove duplicates in a poor communication environment.

2.2 Classification of Deduplication Levels

In the data deduplication environment, file level deduplication and block level deduplication are classified according to the deduplication level.

File-level deduplication is uses a method of hashing the file itself to generate abbreviated hash data, and then comparing the abbreviated hash data with the hash data stored in the storage server. It is possible to search and compare very quickly, so that the data deduplication speed is faster than the block deduplication method. However, due to the characteristics of the hash algorithm, it is recognized as another data even if only one bit of the source data is changed.

In the block-level deduplication, even if the data of the source file is partially changed, since the remaining blocks are the same as the corresponding blocks in the source file, only the remaining blocks are deduplicated. Therefore, the redundancy removal efficiency is higher than the file level deduplication method in which even if only one bit of the file is different, the deduplication is not performed.

3 Secure Data Deduplication

In a cloud storage environment, data deduplication technology creates various security issues in the process of storing data on a remote server. This chapter summarizes these security issues.

3.1 Data Encryption

If the data stored in the data storage is not encrypted, data can't be confidential if the data is leaked by internal or external attack. Therefore, encryption is applied because data must be encrypted to ensure confidentiality. However, it is impossible to perform deduplication through comparison of encrypted data. Therefore, a special encryption technique must be applied for deduplication of encrypted data.

3.2 Dictionary Attack

With CE, even if different users encrypt the same source data, the same ciphertext is generated and can be deduplicated. However, since the same encryption key is always generated in the same source data, there is a problem that the encryption key and the encryption data can be generated by arbitrarily generating data that is not actually owned. An example is an in-house document. If an in-house document has a specific format and the contents of the form are limited, an attacker can list data that can be included in the file of the form and generate data, key, and encrypted data by assigning them one by one. This is called dictionary attack.

3.3 Poison Attack

In the client side deduplication method, the user transmits a list of data to be uploaded to the server, and the server transmits a list of data that is not duplicated to the user to the user. The user transmits only the non-duplicated data to the server based on the received list. However, in this process, when the user uploads the list to be uploaded differently from the actual upload data, the information of the data stored in the server and the actual data become inconsistent. An attack using this is called a poison attack.

4 Security Technology of Secure Data Deduplication

In this chapter, we discuss security techniques for secure data deduplication.

4.1 Convergent Encryption (CE)

CE is a cryptographic scheme proposed by Douceur et al. in 2002 [2]. This technique uses a method of hashing the data source to generate an encryption key, and encrypting the data using this key. This method allows duplicate data to be removed because the same ciphertext is always generated even if different users encrypt the same source data.

4.2 Prepare for Dictionary Attacks

In 2013, Bellare proposed a key generation scheme using RSA OPRF, which is a combination of OPRF designed by Naor in 1997 and RSA blind signature proposed by Bellare, et al. in 2003 and Chaum in 1983 [3–6]. With RSA OPRF, the key can be generated by communicating with the server. In this method, the RSA Blind Signature-based operations are repeatedly performed by the server and the user. This makes it impossible for an attacker to forge an encryption key using a dictionary attack.

Message Locked Encryption (MLE) is a data encryption technology proposed by Bellare et al. In 2013 [7]. This technique proposed a method for generating the same encrypted data from the same data source. The general approach of MLE is to use the data source M to compute $K \leftarrow H(M)$, and use K as the key to generate $C \leftarrow E_K(M)$. MLE includes a total of four encryption technologies.

4.3 Prepare for Poison Attacks

Poison attacks occur when there is a mismatch between the source of secure data and the tags stored in the server [8]. In other words, it occurs when the cipher data generated from the data different from the source of the tag used when the user checks the duplication is uploaded. This causes Loss of Data Source and Damage due to modulated data. Therefore, in order to prevent the poison attack, it is necessary to check whether the uploaded secure data is generated from the same source as the tag used in the check for duplication.

5 Secure Data Deduplication Systems

This chapter examines secure data deduplication systems and compares each system as shown in Table 1.

5.1 DupLESS

DupLESS is a cryptographic data deduplication system proposed by Bellare, et al. In 2013 [3]. The CE or MLE, which is mainly used in the deduplication of encrypted data, is fundamentally vulnerable to dictionary attacks, which has the problem that the attacker can guess the plaintext data only with the encrypted data. DupLESS, on the other hand, uses a key server additionally to provide a secure cryptographic data deduplication system that is resistant to brute force attack.

The user generates the key K in cooperation with the key server via the RSA-OPRF protocol. In the process of communication between two entities, the user can not know the private key which is secret information of the key server, and the key server can't know the message based encryption key which is the secret information generated from the user.

5.2 ClouDedup

Cloudedup is a cryptographic data deduplication system that provides block-level deduplication and data confidentiality proposed by Puzio, et al. It is based on Convergent Encryption, but it is safe against dictionary attacks by introducing a server that performs additional cryptographic functions [9]. Unlike file-based deduplication, block-level deduplication the need to manage the CE keys for each block. In ClouDedup, the metadata manager performs key management and deduplication.

The client divides the file into blocks and encrypts each block with CE. The client only stores the encryption key for the first block.

The Meta Data Manager (MM) stores the encrypted key and block signature values and removes duplicate blocks. In order to organize files from multiple blocks and manage file ownership, we manage file tables, pointer tables, and signature value tables as well as linked lists including block id and file id. The storage server stores the blocks received from the MM and does not play any role in deduplication.

5.3 PerfectDedup

Table 1 Comparison of secure data deduplication systems

	DupLESS	ClouDedup	PerfectDedup
Deduplication level	File level	Block level	Block level
Key generate	Key Server (RSA OPRF)	User	User
Brute force attack resistibility	○	○	○
Dictionary attack resistibility	X	○	○
ETC	-	-	Popularity-based deduplication

○: Offer; X: Not offer;

PerfectDedup is a data deduplication technology developed by Puzio et al. [10]. The existing data deduplication technologies can solve the conflict between data deduplication and data encryption by applying CE. However, such deduplication technologies encode all uploaded data while applying CE, which can increase the computational overhead. Therefore, to address this issue, the popularity of uploaded data can be measured. If the data has a popularity higher than the threshold value, CE could be applied. PerfectDedup to perform data deduplication while providing data confidentiality to end users. Furthermore, PerfectDedup provides decreased accessibility to low-popularity data such as personal information, improves the accessibility of popular data, and thereby helps in reducing the computational overhead.

6 Conclusion

In this paper, we have discussed techniques for performing data deduplication and techniques for safely performing data deduplication. These technologies are not completely redesigned, but have some form of modified security technology applied to the need for data deduplication. While these technologies have achieved their intended objectives, it is difficult to expect high levels of completeness while satisfying both computation and traffic efficiency due to the emergence of additional security threats and the use of additional security technologies to address them. However, the recent demand for privacy and security technologies is expected to increase the demand and supply of data deduplication technology in the future, which is expected to lead to more sophisticated technology research. Future research will analyze new security threats in data deduplication environments and various technologies to solve them. This will discuss whether data deduplication techniques are suitable for real-world environments and how future technology research will be conducted.

Acknowledgments. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2016R1D1A1B03935917) and Barun ICT Research Center at Yonsei University.

References

1. Storer, M.W., et al.: Secure data deduplication. In: Proceedings of the 4th ACM International Workshop on Storage Security and Survivability. ACM (2008)
2. Douceur, J.R., Adya, A., Bolosky, W.J., Simon, P., Theimer, M.: Reclaiming space from duplicate files in a serverless distributed file system. In: 2002 22nd International Conference on Distributed Computing Systems. Proceedings, pp. 617–624. IEEE (2002)
3. Bellare, M., Keelveedhi, S., Ristenpart, T.: DupLESS: server-aided encryption for deduplicated storage. IACR Cryptology ePrint Archive 429 (2013)
4. Naor, M., Reingold, O.: Number-theoretic constructions of efficient pseudo-random functions. J. ACM (JACM) **51**(2), 231–262 (2004)
5. Chaum, D.: Blind signatures for untraceable payments. In: Advances in Cryptology, pp. 199–203. Springer, Boston (1983)
6. Bellare, M., Namprempre, C., Pointcheval, D., Semanko, M.: The one-more-RSA-inversion problems and the security of Chaum's blind signature scheme. J. Cryptol. **16**(3) (2003)
7. Bellare, M., Keelveedhi, S., Ristenpart, T.: Message-locked encryption and secure deduplication. In: Annual International Conference on the Theory and Applications of Cryptographic Techniques, pp. 296–312. Springer, Heidelberg (2013)
8. Kaaniche, N., Laurent, M.: A secure client side deduplication scheme in cloud storage environments. In: 6th International Conference on New Technologies, Mobility and Security, NTMS 2014, pp. 1–7 (2014)
9. Puzio, P., Molva, R., Onen, M., Loureiro, S.: ClouDedup: secure deduplication with encrypted data for cloud storage. In: 2013 IEEE 5th International Conference on Cloud Computing Technology and Science (CloudCom), pp. 363–370. IEEE (2013)
10. Puzio, P., Molva, R., Önen, M., Loureiro, S.: PerfectDedup: secure data deduplication. In: Data Privacy Management, and Security Assurance, pp. 150–166. Springer, Cham

IoT Based Automatic Notification System in Factory Using Walkie-Talkies

Hyunmin Park[1(✉)] and Jaewook Jeon[2]

[1] Department of Digital Media and Communications Engineering, Sungkyunkwan University, Suwon 16419, South Korea
hwayong@skku.edu
[2] School of Information and Communication Engineering, Sungkyunkwan University, Suwon 16419, South Korea
jwjeon@yurim.skku.ac.kr

Abstract. This paper proposes a method of using internet of things (IoT) technology and walkie-talkies to notify factory operators of information about the manufacturing system of the factory. The proposed automatic notification system consists of walkie-talkies, manufacturing IT systems, and a device called the walkie-talkie automatic notification device based on IoT. The role of the walkie-talkie automatic notification device is to act as a gateway that sends information from the manufacturing IT systems to walkie-talkies as voice automatically. It is based on Raspberry Pi 3 hardware. We designed a simple automatic notification system that does not require extra infrastructure costs by using walkie-talkies that are in widespread use at manufacturing sites. This makes it possible to convey information about the machines in a factory or give other instructions to operators easily and efficiently.

Keywords: Notification system · Walkie-talkie · Internet of things (IoT) · Manufacturing system

1 Introduction

Today, internet of things (IoT) technology has improved greatly, due to the development of wireless sensor networks (WSN), embedded systems, and the supply of low-cost sensors [1–3]. This IoT technology is being applied to a variety of fields [4], including home automation [5], healthcare [6], and so on. The manufacturing industry is also evolving through a variety of IoT technologies [7].

For example, IoT technology allows the factory to monitor facility information and sensor data through the manufacturing system [8]. Using the manufacturing system, we can check unexpected situations such as equipment failures, network problems, and shortages of materials. Environmental monitoring is possible too. However, for actual improvements rather than simple monitoring, the information and working instructions should be communicated to the operators at the factory site.

There are several ways to communicate information about the manufacturing system to the operators. First is a method by which the manager can continuously monitor the manufacturing system and contact the operator if necessary. This is the

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 471–476, 2020.
https://doi.org/10.1007/978-981-13-9341-9_81

most basic. But there are difficulties in monitoring the system continuously, and the manager may make errors. The second is that the manufacturing system can broadcast information using an alarm lamp, but there are limitations to the delivery of information, and operators may not be able to see the alarm. Finally, the manufacturing system conveys information automatically to the operators using SMS, email, and so on [9–11]. This method is efficient because it does not require constant monitoring, and information only needs to be provided to those who need it. However, manufacturers that do not have an infrastructure will incur significant costs to build it. Also, some manufacturers may not be able to use these methods for security reasons, because their information must be transferred to an external server.

Considering the problems of the above methods, this paper proposes a new IoT-based automatic notification system in factories using walkie-talkies. The walkie-talkie can be set to its own channel, and only on the same channel can communicate with each other. Thus, if walkie-talkies are set to different channels for each group of workers, it is possible to communicate only information required within each group. And by using walkie-talkies that are already in use in the factory, the cost of installing the system can be reduced. In other words, it is not necessary to install a new infrastructure for the notification system. Additionally, there are no limitation to what can be communicated by voice. In addition, establishing a system that automatically sends information to walkie-talkies relieves the manager of the responsibility of continuously monitoring the manufacturing system. Of course, the use of walkie-talkies is not the best solution for automatic notification on a manufacturing site. However, it is possible to build an effective system using these features.

2 Proposed System

In a factory, manufacturing data is transmitted to the local server generally. Using the data on the server, various manufacturing systems can be built and used, such as energy, machine, environment monitoring, and so on. The proposed automatic notification system is based on these systems.

Figure 1 shows the architecture of the proposed system. It consists of three components: the manufacturing systems, the walkie-talkie automatic notification device (WAND), and the walkie-talkies in the red box shown in Fig. 1. Through the manufacturing systems, we can monitor the status of the manufacturing information, such as machines defects and materials replacement situations. Some information should be communicated to the operators in the factory area for machine inspections or materials replacements. Failure to do so may result in product defects or production interruption. This information could be made in the manufacturing system and can be transmitted automatically. We used walkie-talkies for the devices for automatic communication, and we used the WAND to do this.

The WAND is used to multicast speech messages to walkie-talkies, and it acts as a gateway between manufacturer systems and walkie-talkies. It waits for messages from the manufacturing systems. They communicate by the TCP/IP method. WAND converts the text messages received into speech and multicasts the speech over walkie-talkies using radio frequency (RF) communications.

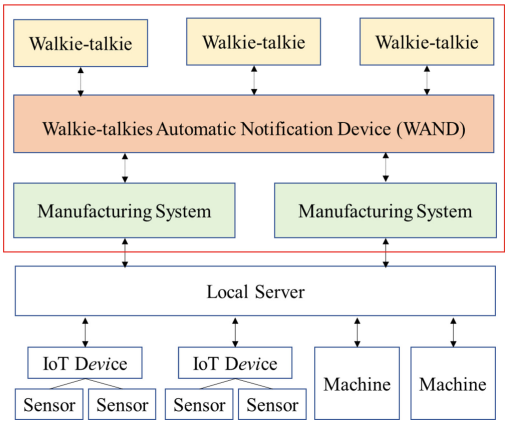


Fig. 1. Architecture of automatic notification system

3 System Implementation

3.1 Walkie-Talkie

A walkie-talkie is a portable, two-way radio transceiver, and it is a half-duplex communication device. Commercial walkie-talkies generally use the ultra-high frequency (UHF) band. The specific frequency for each channel depends on the manufacturer and model. Some can also be programmed to change frequencies for security. Walkie-talkies communicate with others set to the same RF. Tone squelch is an important feature to have in the system too. It is used to solve interference problems in walkie-talkies. When more than one user is on the same channel, tone squelch addresses a subset of all receivers. For this paper, we used the MYT-5800.

3.2 Walkie-Talkie Automatic Notification Device (WAND)

The functions of WAND are divided into three categories. The first is the ability to receive messages from the manufacturing system. For this function, it should create a socket server to receive messages, and it should have memory for storing them. The second function is text-to-speech (TTS). TTS is a computer-based system that is able to read any text. Using the TTS, WAND can generate speech from a text message received from the manufacturing system to send it by walkie-talkies. The last function is RF communication capability to multicast the voice to the walkie-talkies. For these functions of WAND, Raspberry Pi 3 and a UHF band transceiver module named NR-UADTM manufactured by NEOTICS, were used.

The left side of Fig. 2 shows the connection of the hardware. Frequency and tone squelch are set through UART, and GPIO is used to check the transmittable status and change the transceiver mode. The converted speech message is input to NR-UADTM through its audio output. The right side of Fig. 2 shows the connection of the manufactured device referred to.

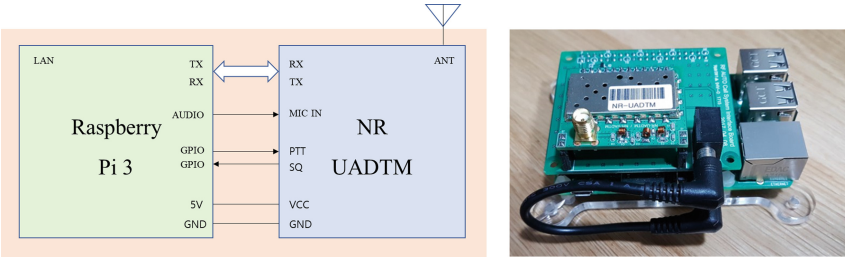


Fig. 2. Hardware connection diagram and actual configuration view of WAND

3.3 Message Transmission Between Manufacturing Systems and WAND

Two threads are executed on WAND. The first thread is for communication with the manufacturing system. Manufacturing systems and WAND communicate using TCP/IP sockets. The left side of Fig. 3 shows the flow chart between the two, where the server is WAND and the client is the system. In the first step, WAND creates a socket server, and it waits for the client to connect. When a situation occurs in which information must be communicated to operators, the manufacturing system creates a client socket, connects to the WAND and sends message. After the WAND receives the message, the received message is put into the queue. Finally, WAND sends the completion message to the client system and goes back for waiting the other client connection. After that, the client terminates the socket communication.

Multiple client systems can be connected and can transmit messages because the WAND is a server. That is, it can be used in conjunction with several systems rather than just one.

3.4 Automatic Notification to Walkie-Talkie

The second thread is used to automatically notify the walkie-talkies. The right side of Fig. 3 shows the flowchart. It starts in receiving mode like a walkie-talkie. When it runs, it checks the data in the queue. If there are one or more data in the queue, it imports the first piece of data. The message from the queue is expressed as, “channel, tone squelch, text message”. In the next step, the channel and tone squelch of WAND are changed using the message. After that, we should check that the SQ signal. A high SQ signal shows the possible condition for multicasting the voice. When the SQ is high, the WAND can be set to the transmitting mode, and the PTT signal. Next, the TTS synthesizer is run using the message. Finally, the WAND multicast the converted speech to the walkie-talkies at the same radio frequency as WAND. After that, WAND goes back to receiving mode and checks data in the queue.

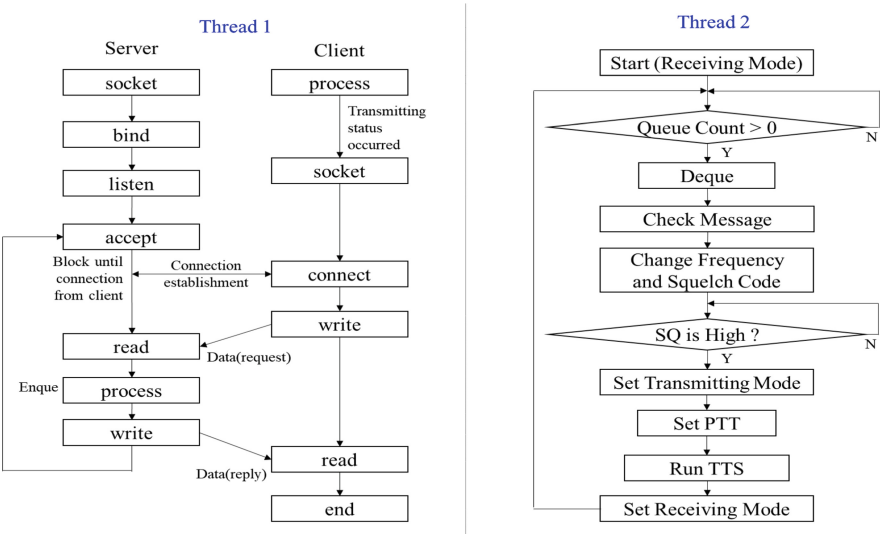


Fig. 3. Flowchart of the WAND threads. Thread 1 is for communicate with the manufacturing systems and thread 2 is for automatically notify the walkie-talkies.

4 System Installation Result

We installed this system in the cable manufacture line of a company that makes medical products. In the cable manufacturing processes, machines test such issues as insulation, opening, and the capacitance of the cables. If any defects are found, the operator cuts the defective section of cable manually. If the operator is unable to act quickly, the number of cables being cut will be higher, and this will cause losses. Therefore, it is necessary to alert the operator to the failure quickly.

Before the system installation, operators got failure information from a lamps and buzzer. However, since operators manage many machines and that area was so noisy, operators sometimes could not recognize the alarms. After the system was installed, operators could become aware of defects in the cable immediately with their walkie-talkies. So defective sections of cable could be removed immediately. As a result, the yields of the process have increased 3.1% over the previous year, from 91.3% to 94.4%. In addition, during the six months in which the system was in use, there were zero defects that went unhandled. The cost of installing the system was also effective. By using existing walkie-talkies, the cost of materials about 200 US dollars (Table 1).

Table 1. System installation result

Contents	Result
Yield improvement	3.1% improved from the previous year
Unhandled defects	0 (for 6 months)
Installation cost	About 200 dollars

5 Conclusion

In this paper, a system was designed and produced that would automatically notify operators of important information at the manufacturing site using walkie-talkies. Through this system, a manager does not need to constantly monitor the manufacturing system in order to notify operators of problems that develop, and there is no need to build costly infrastructure. The WAND was made inexpensively, based on Raspberry Pi, and it was configured with a server so that it can be used in conjunction with multiple manufacturing systems.

Many of the manufacturing sites have changed to digital communication, but walkie-talkies are still used for communication between workers easily and quickly. By using walkie-talkies combined with IoT technologies, we developed an automatic notification system at a low cost. It is expected to contribute to the establishment of smart factories using IoT, such as facility preservation and failure prediction.

References

1. Atzori, L., Iera, A., Morabito, G.: The Internet of Things: a survey. Comput. Netw. **54**(15), 2787–2805 (2010)
2. Aziz, A.A., Sekercioglu, Y.A., Fitzpatrick, P., Ivanovich, M.: A survey on distributed topology control techniques for extending the lifetime of battery powered wireless sensor networks. IEEE Commun. Surv. Tutor. **15**(1), 121–144 (2013)
3. Razzaque, M.A., Milojevic-Jevric, M., Palade, A., Clarke, S.: Middleware for Internet of Things: a survey. IEEE Internet Things J. **3**(1), 70–95 (2016)
4. Da Xu, L., He, W., Li, S.: Internet of Things in industries: a survey. IEEE Trans. Ind. Inform. **10**(4), 2233–2243 (2014)
5. Kelly, S.D.T., Suryadevara, N.K., Mukhopadhyay, S.C.: Towards the implementation of IoT for environmental condition monitoring in homes. IEEE Sens. J. **13**(10), 3846–3853 (2013)
6. Xu, B., Da Xu, L., Cai, H., Xie, C., Hu, J., Bu, F.: Ubiquitous data accessing method in IoT-based information system for emergency medical services. IEEE Trans. Ind. Inform. **10**(2), 1578–1586 (2014)
7. Tandur, D., Gandhi, M., Kour, H., Gore, R.: An IoT infrastructure solution for factories. In: 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 1–4. IEEE (2017)
8. Kwon, Y.J.: IoT-based defect predictive manufacturing systems. In: 2017 International Conference on Information and Communication Technology Convergence (ICTC), pp. 1067–1069. IEEE (2017)
9. Khan, S.F.: Health care monitoring system in Internet of Things (IoT) by using RFID. In: International Conference on Industrial Technology and Management (ICITM), pp. 198–204. IEEE (2017)
10. Jimenez, F., Torres, R.: Building an IoT-aware healthcare monitoring system. In: 2015 34th International Conference of the Chilean Computer Science Society (SCCC), pp. 1–4. IEEE (2015)
11. Promsawat, T., Kummool, S., Pongswatd, S., Julsereewong, A.: Real-time monitoring and reporting alarm system for pH measurement in wet scrubbers. In: 2016 16th International Conference on Control, Automation and Systems (ICCAS), pp. 353–358. IEEE (2016)

A Communication Challenges and Framework for the Control and Monitoring of Autonomous Ship

Kwangil Lee(✉)

Division of Control and Automation Engineering, Korea Maritime and Ocean University, 727 Taejong-ro, Yeongdo-gu, Busan 49112, Republic of Korea
leeki@kmou.ac.kr

Abstract. Recently, an autonomous ship becomes an important issue in the maritime industry. An autonomous ship should be monitored and controlled remotely, i.e. from the shore control center. Since the duty of a human is reduced or eliminated, the safe remote operations of the ship becomes an important issue. For this, the safe and reliable communication from the shore to ship is important. In this paper, we address some communication challenges related with autonomous ship and a propose a message management framework for the autonomous ship which manage the location of the ship and provide the seamless connection among communication parties. It also provides the group communication and QoS.

Keywords: Autonomous ship · Remote control · Location management · Seamless communication · Group communication

1 Introduction

Recently, autonomous ship has been a great attention in the maritime industry. Autonomous ship is a new paradigm in the fourth Industry revolution era. The main motivation of the autonomous ship is the safety, reliability, efficiency and environment. It is expected to make use of various technologies to realize the autonomous ship such as Internet of Things, Big Data, Internet of Services, augmented/virtual reality (AR/VR), Cyber Security, .., etc. [1, 2]. Since it has a great impact on the maritime industry, IMO (International maritime Organization) is also consider the autonomous ship starting from the regulatory scoping exercise. It is expected to cause many changes in the ship design, operation as well as the regulations.

Many issues should be considered for the realization of the autonomous ship. The communication is an important issue to provide the connectivity from a ship to the shore. The shore monitor and control all equipment in a ship so as to take some actions when the emergency situation arises [3, 4]. There are a lot of data bandwidth required. In addition, providing seamless connectivity is also an important issue. In this paper, we address the connectivity issues and propose a communication management system for the autonomous ship.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 477–482, 2020.
https://doi.org/10.1007/978-981-13-9341-9_82

2 Communication Challenges for Autonomous Ship

An autonomous ship requires the communication between a ship to shore. Since the communication intra-structure is not enough, it is always a challenges in the bandwidth and the coverage. In this paper, we briefly address some communication issues for the autonomous ship.

2.1 Communication Links

When a ship has some critical or emergency situation, the ship should be controlled from the shore. So, Always-on-connectivity is essential for the autonomous ship. Even though the communication requirement is not huge in the deep-sea voyage, the situational awareness requires the high volume of the bandwidth [3]. In addition, the communication near the shore increase a lot of communication. Since it requires a complex situational awareness. Since the satellite communication has limited bandwidth, recently 4G/5G communication is also considered within 100 km from the shore [5] (Table 1).

Table 1. Available communication links and bandwidth in a ship

Type	Bandwidth	Category
Inmarsat	432 kbit/s	Satellite
Global Xpress	5–50 Mbit/s	Satellite
Iridium	134 kbit/s	Satellite
VSAT	4 Mbit/s …	Satellite

Autonomous ship consists of several types of communication modules. It is requested that the communication links needs to be handed over seamlessly with the maximum available communication links based on the communication requirement.

2.2 Communication Management Requirement

An autonomous ship requires the collection and the exchange of all equipment and service information of the ship. It also requires the various information from the shore such as weather information and maritime safety information. Also, the shore control center requires the management and control of the ship, equipment and services. The shore control system requires all the information is necessary. The following issues needs to be considered.

Addressing. There is a unique method to define address for ships, devices and services. A ship has several identifiers such as vessel name, IMO number, MMSI and call sign,…, etc. However, not all information is available for all ships. So, it is a common way to define a ship so as to deliver a message to the right destination. In this paper, we use the MMSI number for the identification of a ship.

Location Management. A ship moves around a world. The location management can be achieved in the physical logical location management. The physical location management tracks of the physical GPS (latitude and longitude) information. The logical management is to track of the which communication manager takes care of the delivery of the message.

Our system defines the logical message system which makes use of the physical communication system. In this paper, the location management system only takes care of the logical location management.

QoS. QoS is an important issue for the autonomous ship since the bandwidth in a ship is very limited. A ship requires the communication with many other stakeholders on the shore. When remote control mode with emergency situations or alerts shall be delivered to the shore on time. This might require the utilization of different communication links at the same time.

Scalability. The scalability is also important issue for the autonomous ship. Since the number of ships for the remote control will be increased. In theoretically, the number of ships cannot be limited. The communication system should be enough to control and monitor any number of ships.

3 Communication Management System for an Autonomous Ship

In this paper, we propose a communication management system which is used for an autonomous ship. The communication management system architecture for the autonomous ship is defined in Fig. 1. As illustrated in Fig. 1, the message system consists of shore domains, ship domains and a location server.

3.1 Shore Domain

A shore domain is for a shore side, especially for shore communication manager, i.e. in the shore control center. A shore domain can be connected to a number of ship domains. Since it is possible to have a number of shore control center, it is possible to have multiple shore domains. This means that a ship domain can be connected to multiple shore domains but only one shore domain at a time.

The role of a shore domain is to manage of all equipment and services as well as of all ships itself. The bus manager is to manage all local bus manager of each ship. Each ship is identified uniquely by its MMSI information. Ship manager is to manage all equipment and services. It collects all device and service information from local controllers in each ship. The identification of device and service is performed with the combination of UUID and ship's identity information. When it is identified, the information is registered to the location server.

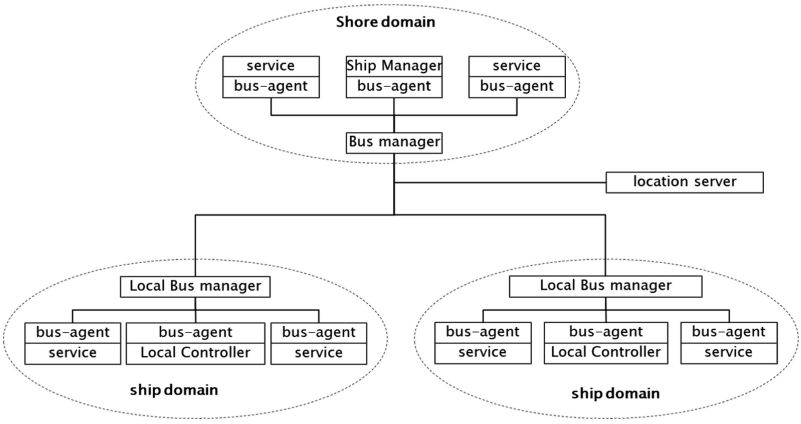


Fig. 1. Structure of the autonomous message management framework.

3.2 Ship Domain

A ship domain represents a ship and connect all components in a ship. A ship requires to be connected to one of shore domain. A ship may choose any of shore domain since all shore domain is connected to Internet.

The choice of the shore domain is based on the QoS such as minimize the delay and response time. Since the real-time information is very important, a shore domain should be selected to minimize the delay and response time.

As illustrated in Fig. 1, a shore domain also consists of bus manager and bus agent. A bus manager is connected to local controller and a bus agent is connected to each service. A bus manager manages the connectivity and the communication with all bus agents. A bus manager also acts as a gateway for a ship which provides a connection with shore domain. In other words, any service wants to communicate with a shore, it first forwards the messages to the local bus manager, then the local bus manager relays the message to the Bus Manager.

A local controller manages all service within a ship. This means that all service needs to be registered and un-registered to the local controller before it uses. The local controller reports the status of each service to the ship manager.

3.3 Location Server

A location server is to have a mapping table for shore domain and ship domain. In our system, there is only one location server.

When a ship domain connects to a shore domain, a shore domain registers the registered information to the location server. The location server keeps a record of the ship identification information and current shore domain manager information. A location server is required to have up-to-date location information of each ship.

There is only one location server in our system. However, the location server can be implemented in a distributed manner. In this case, all location information needs to be

shared among all location servers. The issue for how to synchronize the location information is not the scope of this paper.

4 Autonomous Message Management Framework

The detailed message management framework is illustrated in Fig. 2. It consists of message handler, massage broker and massage manger.

4.1 Message Handler

A message handler is to send and to receive messages. This means that the message handler has a responsibility of the message packetization and de-packetization. Also, it performs protocol conversion for the inter-operability since each device and equipment may uses a different message format,

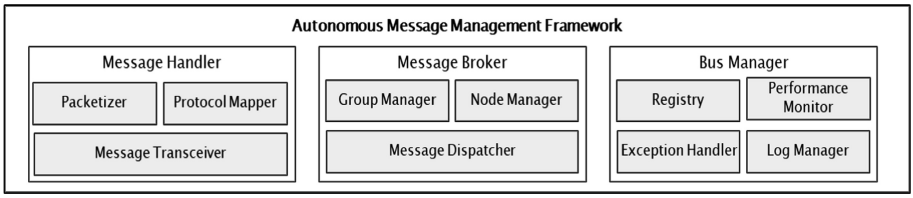


Fig. 2. Structure of the autonomous message management framework.

4.2 Message Broker

The message broker is to manage a connectivity of each node and manage the group communication. Node manager manages the connectivity and the status of the communication. It also manages the group membership information. When a message is destined to a group, it performs the group communication by duplicating a message to each destination.

4.3 Message Manager

The message manager is to perform the manage the message delivery and management of the message system in a ship. The message manager manages a list of connected message agents using the registration and initialization process. all It also performs the performance monitoring and log management. Finally, it provides exception handling procedures whenever any exceptions occur.

5 Conclusion

Recently, autonomous ship becomes an important issue in the maritime sector. An autonomous ship is considered to be monitored and controlled remotely, i.e. from shore control center. Since the duty of a human is reduced or eliminated, the safe and reliable control of the ship from the shore becomes an important issue. In this perspective, the safe and reliable communication from the shore to ship is important. In this paper, we address some communication issues related with autonomous ship and a propose a framework for the message system which manage the location and provide the seamless connection among communication parties. The safe and secure communication infrastructure is left for the future study.

References

1. Rødseth, Ø.J., Lee, K.: A taxonomy of autonomous merchant ship. In: International Symposium on Maritime Technology 2017, Busan (2017)
2. Lee, K.: A remote control and monitoring based on common maritime data structure for autonomous ship. In: International Symposium on Maritime Technology 2017, Busan (2017)
3. Rødseth, Ø.J., May, P., Ehrlich, Kvamstad, B., Porathe, T.: Communication architecture for an unmanned merchant ship. In: IEEE Oceans 2013. IEEE Press, New York (2013)
4. Höyhtyä, M., Huusko, J., Kiviranta, M., Solberg, K., Rokka, J.: Connectivity for autonomous ships: architecture, use cases and research challenges. In: 8th International Conference on ICT Convergence, October 2017 (2017)
5. Poikonen, J.: Ship communication challenges for autonomous and remote-controlled vessels. European Maritime Day (2016)

High-Speed Horizon Line Detection Algorithm Using Dual Hough Algorithm for Marine Image

Hyeoncheol Shin and Kwangil Lee(✉)

Division of Control and Automation Engineering, Korea Maritime and Ocean University, 727 Taejong-Ro, Yeongdo-Gu, Busan 49112, Republic of Korea
{guscjf365,leeki}@kmou.ac.kr

Abstract. Horizon detection is the first step for analyzing video-based maritime situations on autonomous vessels. The detected horizontal line information is used for various aspects such as distance estimation, image registration, and the region of interest detection. Recently, a method of improving the accuracy of horizontal line detection by using multi-scale image and radon transform has been proposed, but there is a problem that detection time is long. In this paper, we propose a method to accurately detect the horizon at high speed for the marine image. We set a color-based region of interest, which is called ROI and uses Duale Hough Transform which improves the speed of computation and accuracy. In the experimental analysis, the proposed method shows similar accuracy for the horizontal line detection with MuSCoWERT and shows the similar processing time with Hough transform algorithm. The experimental results illustrate that both the processing time and accuracy can be improved together up to optimal with the combination of ROI and the Dual-Hough transform.

Keywords: Horizon detection · Hough transform · ROI · Image processing

1 Introduction

With the advent of autonomous vessels, the maritime transport industry is undergoing dramatic changes into the autonomous shipping. In addition, the importance of safety has increased, and methods for safe operations of ships in various fields have been proposed. The situational awareness with sensor fusion becomes more important for the safe operation of the autonomous ship so as to avoid collisions with other ships or other objects. For the situational awareness, camera with the sensor fusion technology becomes more important for the realization of the autonomous ship which detects and identifies surrounding objects by combining various other sensors information includings radar, AIS and Lider. The issues is how the camera can detect object, classifies the objects accurately with the minimum processing time.

In the marine image processing, the horizontal line is the most important factor. It is very important to deduce or utilize information through the horizon because it is difficult to obtain data that is different from land. For example, you can measure the distance from the object by measuring the distance using the horizontal line at sea.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 483–488, 2020.
https://doi.org/10.1007/978-981-13-9341-9_83

In addition, when the camera mounted on the ship is shaken, the images can be matched. Since the horizon is used as an index to provide additional information in the marine environment, detecting the horizon is the first step in the process of analyzing the maritime image.

When detecting a horizontal line, it is detected based on a linear characteristic such as Hough transform or Radon transform [1]. In addition, the color difference between the sky and the sea is analyzed in terms of pixels, and the boundary is statistically detected [2]. Using the advantages of these methods can detect the horizontal line, but there are some problems.

First, detection using only linear features interferes with the detection of horizontal lines due to noise other than the linear characteristics of the horizon, such as a long ship, a wave, or a long wake caused by the passing of a ship. To overcome this problem, a method of extracting the linear features after removing the ambient noise using a multi-scale edge detector has been proposed [3]. However, it takes a long time and it is difficult to apply in real time.

In this paper, we propose a method to detect horizon in real time by detecting horizon at high speed with the optimal line detection accuracy.

2 ROI and Dual Hough Transform Algorithm

The flow chart of the proposed method is shown in the following Fig. 1. In the maritime image, the horizon, which is the boundary between the sky and the sea, is characterized by a large color change. Through this, we set the area of interest and improved the processing speed. Then, the edge is detected from the average image of the images obtained by performing the median operation at various scales. Detecting an edge in an average multi-scale image can solve the instability of edge detection and improve the processing speed. Finally, a horizontal line is finally detected by using a Dual-Hough transformation as a method of detecting a linear feature.

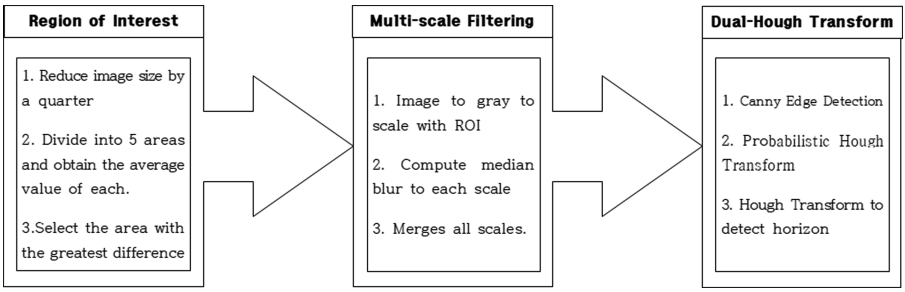


Fig. 1. Flowchart of the proposed method

2.1 Region of Interest

First, we reduced the image to 25%, which improved the computation speed. When the color change is detected by reducing the original image, the difference between the original image and the original image is insignificant, so that the operation speed can be improved. In the reduced image, each frame is divided into five regions. Select the part with the most color difference in five areas and detect the horizon in that area. The method of detecting the spot where the color difference is greatest is to perform the first derivative on the basis of the vertical components to detect the place where the most difference occurs. As a result, it is possible to exclude the linear characteristic which occurs in the noise such as waves, and the accuracy and the calculation speed can be improved (Fig. 2).

Fig. 2. The process of setting the region of interest

2.2 Median Multi-scale Filtering

Use a median multiscale filter on the set region of interest to eliminate ambient noise. Smoothing is performed by performing a median multiscale filter of noise with linear characteristics, such as a long ship near the horizon.

Perform the median operation in 10 steps. The image processed by each scale is overlapped with the image of the previous scale to finally generate noise-free image. This strengthens the components with strong intensities, such as horizontal lines, and reduces the components with weak intensities to remove noise (Fig. 3).

Fig. 3. The process of median multi-scale filtering

2.3 Double Hough Transform

Since the horizontal line is a straight line, the linear characteristic is the strongest place. So, the component with the strongest linearity is selected as the horizontal line through the Hough transform.

In this paper, we first select a random pixel value by probabilistic Hough transform and detect the line segments with a certain length. The probabilistic Hough transform detects the line segments with relatively short linear features, and thus can compensate for the obstruction of the horizon due to the ship or suspended matter. Due to the

limited range of the set region of interest, it is possible to reduce the accuracy error by reducing the noise having a linear characteristic such as a wave. If you perform the Hough transform based on the straight line drawn by probabilistic Hough transform, you can finally detect the correct horizontal line (Fig. 4).

Fig. 4. The process of double Hough transform

3 Experimental Results

Figure 5 illustrates the measurement method for comparison of the horizon which is used in this paper. The Y parameter means the vertical distance of the horizontal line from the center point of the top x-axis end. The α parameter is the angle formed by the normal to the horizontal line that points to the sea. That is, the objective is to express the numerical value of the actual horizon by expressing the straight line as the value of Y and α, expressing the distance from the center of the X-axis to the horizon and the angle α thereto, and comparing with the measured value. For the performance measurement, 11 full HD images of 2,772 frames provided by Singapore-Marine Dataset [4] were used. The parameters Y and α represent position and angular error. We used the percentage that sorting position and angle error in ascending order in Singapore-Maritime datasets.

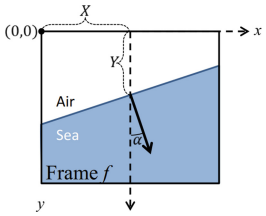


Fig. 5. Measurement method for comparing detected horizon line [3]

Table 1. Values of mean processing time per frame (seconds) for different methods.

Method	**Proposed**	MuSCoWERT [1]	ENIW [1]	Hough [1]
Time (Seconds)	**0.32**	9.2	∼hours	0.3

The experimental results are illustrated in Tables 1 and 2. Table 1 shows the processing time and Table 2 shows the experimental results of position and angular error in each method. MuSCoWERT has the best performance but the detection speed is low (9.2 s). Also, the processing speed is the fastest using a single Hough transform (0.3 s), but the accuracy is the lowest. Compared to that, the proposed method is

accurate enough that the difference in accuracy is insignificant when compared to MuSCoWERT, and the speed is also fast enough to make only 0.02 s difference with a single Hough transform (Fig. 6).

Table 2. Statistics of errors in Y and α, i.e., $|Y - Y_{GT}|/|\alpha - \alpha_{GT}|$, for different methods.

Method	Percentile position error(pixel)/angle error(degree)		
	25^{th}	50^{th}	95^{th}
Proposed	**0.82/0.11**	**1.82/0.30**	**16.23/0.99**
MuSCoWERT [1]	0.54/0.06	1.49/0.25	8.17/0.88
ENIW [1]	1.82/0.30	117.81/1.10	224.48/4.88
Hough [1]	16.23/0.99	221.02/1.89	384.01/7.83

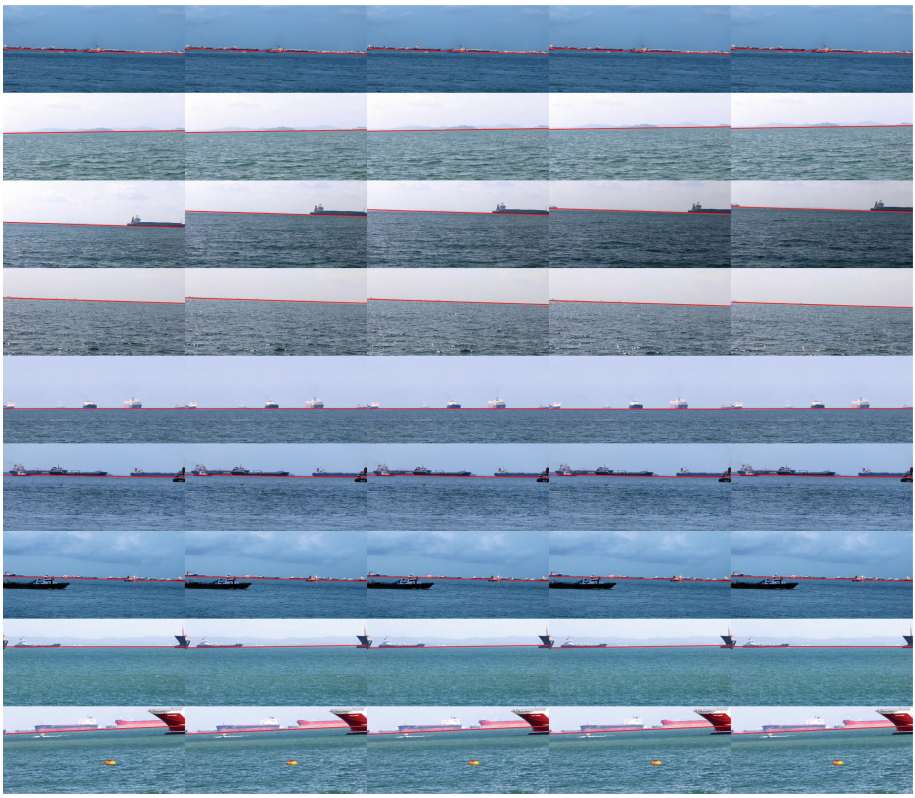


Fig. 6. The red line indicates the horizon detected by proposed method.

4 Conclusion

In this paper, we propose a method to detect the horizontal line at high speed using the detection of the region of interest and the Dual-Hough transform using the characteristics of the horizon in the marine image. Recently, a method of improving the accuracy of horizontal line detection using multi-scale image and radon conversion has been proposed [1], but it has a problem that detection time is long. Experiments on public datasets confirm that the proposed method is accurate and can detect horizon at high speed. Applying the proposed method in this paper, the detection and classification of image - based objects in smart vessels remains a subject of future research.

References

1. Prasad, D.K., Rajan, D., Rachmawati, L., Rajabally, E., Quek, C.: MuSCoWERT: multi-scale consistence of weighted edge radon transform for horizon detection in maritime images. J. Opt. Soc. Am. A **33**, 2491 (2016)
2. Zafarifar, B., Weda, H., de With, P.H.N.: Horizon detection based on sky-color and edge features. In: 2008 Proceedings of SPIE 6822, Visual Communications and Image Processing, p. 682220, 28 January 2008. https://doi.org/10.1117/12.766689
3. Prasad, D.K., Rajan, D., Prasath, C.K., et al.: MSCM-LiFe: multi-scale cross modal linear feature for horizon detection in maritime images. In: IEEE Region 10 Conference, Singapore, Singapore, November 2016, pp. 1366–1370. https://doi.org/10.1109/tencon.2016.7848237
4. Singapore Maritime Dataset. https://sites.google.com/site/dilipprasad/home/singarpore-maritime-dataset

Service Integration Methodology for Convergence Service in Science and Technology

YunHee Kang[1(✉)], R. Young Chul Kim[2], and HeeSeok Choi[3]

[1] Division of Information and Communication, Baekseok University, 115 Anseo-dong, Cheonan, Chungnam 330-704, Korea
yhkang@bu.ac.kr
[2] Department of Software and Communications Engineering, Hongik University, Seoul 30016, South Korea
bob@hongik.ac.kr
[3] Convergence Service Center, KISTI, 245 Daehak-ro, Yuseong-gu, Daejeon 34141, South Korea
choihs@kisti.re.kr

Abstract. To facilitate the open science in research lifecycle, reconstructing individual services into modular and independent function blocks that can be reusable and interoperable. Over the years, more REST based web services have been deployed in the enterprise application that is integrated with services from different sources with existing functionality over HTTP. This results in an increasing interoperability and supporting seamless integration compared to traditional web services. Despite this trend, there are still no standards or guidelines about how to develop a RESTful web service for S&T knowledge infrastructure. In this paper we introduce the methodology for service integration for modelling conceptual architecture and understand the practices on the portal composed of services of convergence of S&T. In the research lifecycle, S&T services are used to support resource activities from developing ideas to sharing to proliferating research results.

Keywords: Service integration · Convergence service · Knowledge infrastructure · Web services · Research lifecycle

1 Introduction

In a new science research way represented as data intensive paradigm, open science drives an innovative approach to the scientific process based on cooperative work and new ways of convergence knowledge by using digital technologies and new collaborative tools. It is about extending the principles of openness to the entire research lifecycle. It can reduce the barriers for sharing any kind of output, resources, methods or tools, at any stage of the research process [1, 2].

To facilitate the open science in research lifecycle, reconstructing individual services into modular and independent function blocks that can be reusable and interoperable. The vision of a service integration for the convergence service of S&T

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 489–493, 2020.
https://doi.org/10.1007/978-981-13-9341-9_84

(Science and Technology) knowledge infrastructure is to seamlessly access computing resources and the research results including artifacts digitally enabled scholars, researchers, and engineers participating in multidisciplinary collaborations in the open science way. The integrated portal with convergence services can provide a systematic and interconnected S&T knowledge infrastructure including NTIS, NDSL, COREEN, etc.

Over the years, more REST based web services have been deployed in the enterprise application that is integrated with services from different sources with existing functionality over HTTP. This results in an increasing interoperability and supporting seamless integration compared to traditional web services. Despite this trend, there are still no standards or guidelines about how to develop a RESTful web service for S&T knowledge infrastructure. In this paper we introduce the methodology for service integration for modelling conceptual architecture and understand the practices on the portal composed of services of convergence of S&T. In the research lifecycle, S&T services are used to support resource activities from developing ideas to sharing to proliferating research results.

2 Related Works

Increasingly, scientific breakthroughs represented as Fourth Paradigm of science have been driving by on-demanding and federated computing resources that help researchers manipulate and explore massive data [2, 4]. But upcoming infrastructure technologies such as cloud computing and multicore processors provide the specific solutions in pre-processing and post-processing the large dataset as well as their tradition computing resources and tools. For example, particle physics has been used to analyze the world's largest datasets by using Hadoop distribution and Spark [2, 3, 5]. Nation-wide research involves many different stages and activities from developing ideas to proliferating those research results.

Service-oriented architecture (SOA) presents a fundamental shift in dealing with the difficulties of building an enterprise information system composed of diverse services involved in research lifecycle for many S&T research areas [6, 7]. It promotes loose coupling between software components so that they can be reused. Applications in SOA are built based on independent services. As the extension of SOA, the microservice architectural style is an approach to developing a single application as a suite of small services, each running in its own process and communicating with lightweight mechanisms, often an HTTP resource API. Cloud computing and SOA can be pursued that cloud computing's platform and storage service offerings can provide a value-added supporting for SOA efforts as an integration technology.

3 The Proposed Service Integration Methodology

3.1 Integration Modelling for Convergence Services

In the enterprise modelling for a convergence service in S&T knowledge infrastructure, we propose the service concept plays a central role. A service is defined as a unit of functionality that an entity (such as a system, organization, or department) makes available to its environment, and which has some value for certain entities in the service users.

Service orientation approach typically leads to a layered view of enterprise architecture models, where the service concept is one of the main linking points between the different layers. Service layers with service made available to higher layers are interleaved with implementation layers that realize the services. Within a layer, there may be internal services, services of supporting applications that are used by the end-user applications. In Fig. 1, each service has specific user and user groups. Two or more services can be integrated into convergence service as business services in integrated portal.

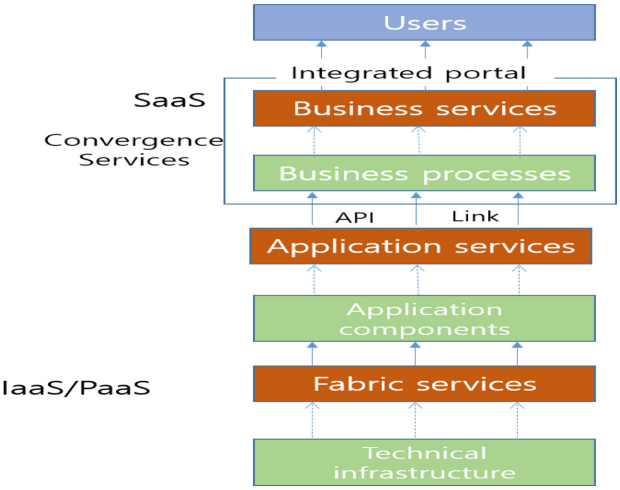


Fig. 1. Layer view for convergence service

With these observations, we decide that modelling enterprise architecture should focus on inter application-domain relation. It provides a metadata for integration by allowing the creation of models that showing how the implementation layers make use of services of other layers located in lower layers, and realization relations, showing how services are realized in an implementation layer. Hierarchical layers are used to separate the business of users, the service, and the gateway. The interaction between layers will most likely be a client-server relation. There is no enforced coherence of central shared data. Working on integration emphasizes the interoperability of services in distributed information systems.

3.2 The Design Guideline for a Convergence Service

Researchers, funding agencies, and publishers are increasingly concerned with matters of scientific reproducibility. Especially the need for integration of services is getting important to users who is working for convergence studies in the research community. New information architectures enable new approaches to publishing and accessing valuable data and programs. SOA allows for the reuse of existing assets where new services can be created from an existing IT infrastructure of systems. In other words, it enables businesses to leverage existing investments by allowing them to reuse existing applications and promises interoperability between heterogeneous applications and technologies.

In this reason, SOA is a good way to handle the complexity of enterprise software environment has been increased. We can build the convergence service by referring the specification of well-designed services and putting together them in bottom-up manner. We apply the cloud-based delivery model for supporting standards interface like REST via light-weight client. In a SOA environment, a service is composed with the operations that the service supports; defining the protocol used to invoke those operations over the Internet; and operating a server to process incoming requests.

(1) **(Loose coupling)** To meet the loose coupling requirement of an enterprise application, the application should communicate a data service mediator with different access mechanisms, synchronous RPC and pub/sub based asynchronous communication. To reduce coupling between applications, we use callback in which the subscriber is asked to supply a reference to a function with a certain signature, which is called later. The publisher of the event knows nothing about the module that is called, so there's no dependency. This communication mechanism is also reduced dependencies between two entities.
(2) **(Interoperability)** To meet the interoperability needs of applications, we consider data is formatted as XML. To reformat the data from data service mediator to the end-user application, we design the proxy to transform the result of a query for different formats.
(3) **(Seamless integration)** As data sizes increase, and researchers rely on a distributed large scale of dataset and publication, data storage, and compute resources, these challenges are likely to become barriers. Hence it is important to support stakeholders for a scientific research that we construct a software architecture for integrating seamlessly separate service in KISTI. We consider applying LISI model for measuring the integration level.
(4) **(Flexible service pool)** The complete set of services for all resource types forms a service pool which provides uniform access to data stored in the underlying information system.
(5) **(Performance)** Handing component interactions can be the dominant factor in user-perceived performance and network efficiency. To manage web services, it needs to monitor quality of them periodically or aperiodically when changing their status like availability, performance and security policy. To meet the performance needs of applications, we consider an optimized binary data format and a data caching.

(6) **(Scalability)** simplifies component implementation, reduces the complexity of connector semantics, improves the effectiveness of performance tuning, allowing to support large numbers of components and interactions among components.
(7) **(Cross-domain accessibility)** it allows to provision users and accounts to REST-based endpoints that the CA API Gateway supports. Identity and Access management (IAM) describes the management of individual identities, their authentication, authorization, roles and privileges within or across system and enterprise boundaries It needs for enterprise or cross-network identity management, which reduces the development time for web service connectors.

4 Conclusion

In forth paradigm era, scientific methods to analyze and organize data are required to handle large-scale datasets for scientific discovery by on-demanding computing resources. Hence, these methods also need to support effective S&T knowledge infrastructure in research life cycle. We applied the practices on the portal composed of services and identified and analyzed requirements for integrating diverse individual services for the logical design of RESTful web services in the S&T knowledge infrastructure.

Acknowledgements. This Research was supported by the Korea Institute of Science and Technology Information (KISTI) in 2018.

References

1. David, P.A.: Understanding the emergence of 'open science' institutions: functionalist economics in historical context. Ind. Corp. Change **13**(4), 571–589 (2004). https://doi.org/10.1093/icc/dth023
2. Gray, J., Liu, D., Nieto-Santisteban, M., Szalay, A., Heber, G., DeWitt, D.: Scientific data management in the coming decade. SIGMOD Rec. **34**, 34–41 (2005)
3. Han, Y.: Bioworks: a workflow for automation of bioinformatics analysis processes. Int. J. Bio-Sci. Bio-Technol. **3**(4), 59–68 (2011)
4. Hey, A.J., Tansley, S., Tolle, K.M., et al.: The Fourth Paradigm: Data-Intensive Scientific Discovery. Microsoft Research Redmond, WA (2009)
5. Sehrish, S., Kowalkowski, J., Paterno, M.F.: Spark and HPC for high energy physics data analyses. In: 2017 IEEE International Parallel and Distributed Processing Symposium Workshops, IPDPS Workshops 2017, Orlando, Buena Vista, FL, USA, 29 May–2 June 2017, pp. 1048–1057 (2017)
6. Erl, T.: Service-Oriented Architecture: Concepts, Technology, and Design. Prentice Hall PTR, Upper Saddle River (2005)
7. Spiess, P., Karnouskos, S., Guinard, D., Savio, D., Baecker, O., Moreira Sá de Souza, L., Trifa, V.: SOA-based integration of the Internet of Things in enterprise services. In: IEEE International Conference on Web Services, ICWS 2009, Los Angeles, CA, USA, pp. 968–975 (2009)

The SA Management Scheme Based on Blockchain for Convergence Service in S&T

YunHee Kang[1(✉)], R. Young Chul Kim[2], and HeeSeok Choi[3]

[1] Division of Information and Communication, Baekseok University, 115 Anseo-dong, Cheonan, Chungnam 330-704, Korea
yhkang@bu.ac.kr
[2] Department of Software and Communications Engineering, Hongik University, Seoul 30016, South Korea
bob@selab.hongik.ac.kr
[3] Convergence Service Center, KISTI, 245 Daehak-ro, Yuseong-gu, Daejeon 34141, South Korea
choihs@kisti.re.kr

Abstract. A convergence service can be built by referring the specification of well-designed services and putting together them in bottom-up manner. In the service system for delivering an S&T convergence service, a service system needs to define a service agreement (SA) and to establish to support continuously and dependably the SA. In the paper we propose a SA management scheme that is used for storing the SA values co-created into the blockchain. The proposed scheme can provide a guideline for designing the S&T service system on SOA framework with a blockchain for providing information about SA as part of service metadata.

Keywords: SOA · Service system · Service agreement (SA) · SA management scheme · Blockchain

1 Introduction

A convergence service can be built by referring the specification of well-designed services, and putting together them in bottom-up manner. Hence, we introduce a service integration model to deliver a service for convergence S&T knowledge infrastructure in a Service-oriented architecture (SOA) based on service platform so that it can easily meet the service agreement (SA) for them. The SA is a contract between a service provider and a service consumer [4, 5]. The contract includes the non-functional requirements of the service met between them.

In the service system for delivering an S&T convergence service, the service system needs to define SA, and to establish to support continuously and dependably the service agreement. Therefore, it is important to keep automatically the SA in the S&T knowledge infrastructure with services owned and managed by internal or external entities.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 494–498, 2020.
https://doi.org/10.1007/978-981-13-9341-9_85

In the paper we propose the SA management scheme that is used for storing the SA values co-created into the blockchain. The scheme contributes to tackle its trustiness issue with a conflict among multi-parties by a smart contract between a service provider and a service consumer. The main contributions are listed as follows:

- design the core permissioned blockchain, and define its business models for the SA management
- design the metadata for services maintained by a distributed ledger for maintaining information about SA with transparency and traceability without interference of intermediary.

2 Related Works

SOA is a strategic framework technology for designing and developing software in the form of interoperable service [5, 6]. A service is defined as a unit of functionality that an entity (such as a system, organization, or department) makes available to its environment which has some value for certain entities in service users. SOA is a way to handle the growth of complexity of an enterprise software environment for providing a convergence service both a technical and a functional side.

Over the past several years, the blockchain technology has been driving a breakthrough in many domains including a p2p based distributed system, and introducing a business use case in diverse industries, including finance, real estate, healthcare, and transactive energy. This innovation rooted from the promise of Bitcoin [1, 2], which is a p2p electronic cash operated to solve the double-spending problem without any intervention of a trusted intermediary. Bitcoin is the first application of blockchain. Blockchain provides specific properties (such as decentralization, transparency, and immutability) that have allowed Bitcoin to become a viable platform for "trustless" transactions by well-known techniques including cryptography, digital signature and consensus mechanism.

On the other hand, a blockchain platform, Ethereum, extended the capabilities of the Bitcoin blockchain by adding a new feature, smart contract [3]. A smart contract is a program that directly controls the exchanges or redistributions of digital assets between two or more members according to certain rules or conditions confirmed between involved members. It provides a single version of truth where participants use digitally signed blocks and consensus algorithms to alter transactions and documents using further blocks.

A blockchain is a decentralized system platform that maintains chronically a list of transactions grouped into blocks constructed via a consensus mechanism to keep information consistently. Figure 1 shows a block consists of a set of four transactions, t0 to t3, where t0 is happened before t1 and t3 occurs lastly in the causal relationship. As shown in Fig. 1, SA represented as a digital asset is inputted to a hash function, H. The output of the hash function H is the hashed SA, SA′ that is used as a part of the transaction t2.

One block can be added to the blockchain at a time. Each block is verified to ensure that it follows in sequence from the previous block. All SA transaction records are kept

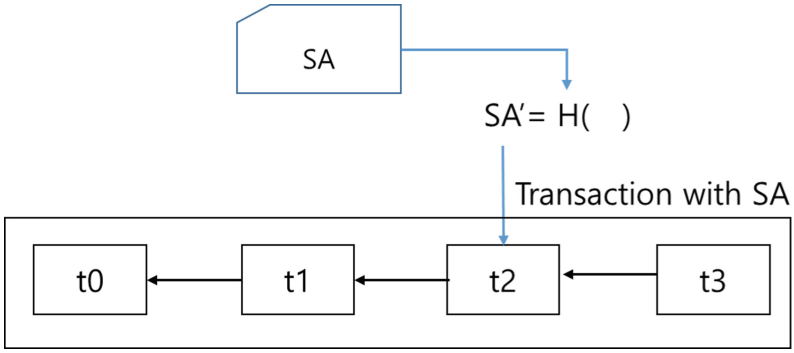


Fig. 1. A set of transactions in a block

in the blockchain and are shared with peers, some members of a service system who have joined in a blockchain network. This decentralized process ensures properties of transparency, data integrity, and robustness

3 The SA Management Scheme Based on Blockchain

In this section, we design the service metadata for maintaining SA in a distributed ledger of a blockchain and the software architecture for blockchain based SA management scheme. The basic idea of our work starts from designing a SOA framework to enable the SA based requirements in publishing and discovery of services. The proposed scheme can provide a guideline for designing the S&T service system SOA framework with a blockchain for providing information about SA as part of service metadata.

By design, the blockchain is a decentralize technology. Blockchain technology has built-in robustness. Hence it may make types of record keeping like SA. By storing blocks of information that are identical across its network, the blockchain cannot be controlled by any single entity and has no single point of failure.

In a SOA structure, a service is a unit of work done by a service provider to achieve a useful result of performing some activities for a service consumer by a service request. Each service implements a specific business function and is made available such that the service can be accessed without knowledge of its underlying implementation. In a software context, services are typically thought of as methods, components, or building blocks of a larger automated system. In the service deliver process, there are three components such as service provider, service consumer and service broker that interact for delivering a specific server in asynchronous manner. There is the essential process of describing the service endpoint as a service metadata so that client can understand how to use the service.

The SA as a legal contract implies that each party to the agreement commits to providing the other party or parties with something they need or want. It underpins the joint understanding between a service provider and a service consumer of what to expect from their mutual relationship. As shown Fig. 2, a service is delivered through the operation of a service system controlled by a service contract.

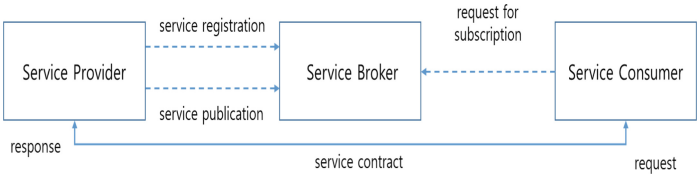


Fig. 2. Service deliver process

Table 1 shows the basic definition of the components of SA blockchain. SA is represented as a digital asset concerned in the blockchain. The participants of the blockchain, who the stakeholders of SA are, include a service consumer, a service provider and a service broker. In the SA blockchain, the transaction can be considered as three operations related with SA. To access the permissioned blockchain, Certificate Authority (CA) and X.509 is used for defining the format of public key certificates. When a certificate is signed by a trusted certificate authority, or validated by other means, someone holding that certificate can rely on the public key it contains to establish secure communications with another party, or validate documents digitally signed by the corresponding private key.

Table 1. The basic definition of the components of SA blockchain.

	Component	SA blockchain
1	Digital asset	Service agreement
2	Participant	Service provider
		Service consumer
		Service broker
3	Transaction	Read the current value of SA
		Read the defined SA value
		Write the value of SA updated
4	Access control	Read by Service provider
		Read by Service consumer
		Write by Service broker
5	ID management	CA and X.509 public/private key

The participants in the blockchain network should authenticate to add further changes in the transactions. It leverages a business ecosystem that represents several integrated business transactions for handling service agreement. The ledger is encryptions and available to all participants. A trusted intermediary is required for transactions related with SA between the service systems who wish to manage the quality of services. Blockchains allow us to have a distributed peer-to-peer network where non-trusting members, both a service provider and a service consumer can interact with each other without a trusted intermediary in a verifiable manner. Figure 3 shows the SA management scheme based on blockchain.

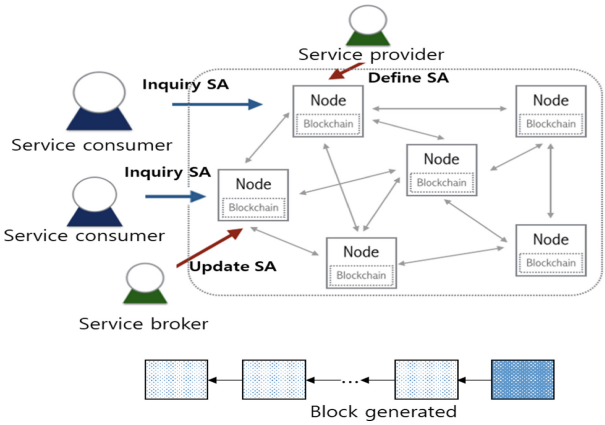


Fig. 3. The SA management scheme based on blockchain

4 Conclusion

In this paper, the SA parameters have been managed by the SA management scheme based on blockchain for convergence service in S&T. This scheme is used for finding the service during services discovery process. The designed scheme is used as a reference architecture for construction of a prototype for verifying the integrity of SA ledger based on block and testing feasibility of the S&T knowledge infrastructure defined.

Acknowledgements.. This Research was supported by the Korea Institute of Science and Technology Information (KISTI) in 2018.

References

1. Nakamoto, S.: Bitcoin: A Peer-to-Peer Electronic Cash System (2009)
2. Swan, M.: Blockchain: Blueprint for a New Economy. O'Reilly Media, Sebastopol (2015)
3. Szabo, N.: Smart contracts: formalizing and securing relationships on public networks. First Monday, vol. 2, no. 9 (1997). http://firstmonday.org/article/view/548/46
4. Patrick Th, E., Rachid, G., Joe, S.: Distributed asynchronous collections: abstractions for publish/subscribe interaction. In: Proceedings of the 14th European Conference on Object-Oriented Programming, pp. 252–276. Springer (2000)
5. François, A.R.J.: Software Architecture for Computer Vision: Beyond Pipes and Filters. Prentice Hall (2003)
6. Erl, T.: Service-Oriented Architecture: Concepts, Technology, and Design. Prentice Hall PTR, Upper Saddle River (2005)

Energy-Based Routing with Rendezvous for DTN-Based UAV Networks

Taehyuk Kim, Euiri An, Wooyeob Lee, and Inwhee Joe(✉)

Department of Computer Software, Hanyang University, Seoul, Korea
{thkim92, reegoon, matiasl2, iwjoe}@hanyang.ac.kr

Abstract. Delay/Disruption Tolerant Networking (DTN) is an approach suitable for networks where end-to-end connections are not guaranteed. In this paper, the energy-based routing with rendezvous is proposed to improve delivery probability of important messages in DTN-based UAV networks. The main idea for achieving this goal is to introduce the rendezvous mode for DTN-based UAV networks, whether to hover or to follow. Simulation was performed using Opportunistic Network Environment (ONE). The simulation results showed that the proposed routing scheme with rendezvous has a larger delivery probability for important messages despite their sizes, compared to the previous routing scheme without rendezvous.

Keywords: DTN · Energy-based routing · UAV

1 Introduction

TCP/IP or ad-hoc networks transmit messages using pre-defined routing paths and enable stable data transmission in end-to-end connected state. However, these existing protocols are not suitable for UAV networks where network disruption is frequent [1].

DTN is an approach suitable for the environment where network disruption occurs frequently. It uses store-and-forward message delivery mechanism to ensure continuous end-to-end connectivity. The bundle layer of DTN store messages to be forwarded in the buffer and forward them on contact with other nodes.

The delivery probability of important messages is very critical for UAV application such as monitoring systems that must respond quickly to emergencies. However, mobile nodes such as UAV have shorter communication time than fixed nodes. This means that UAVs should meet with as many different UAVs as possible. Also, the energy of UAVs is a factor that can have a significant impact on delivery probability of important messages. Because UAVs can return to the waypoint without forwarding important messages due to lack of energy. However, conventional routing schemes used in DTN do not consider the energy of nodes.

To solve these problems, this paper proposes energy-based routing with rendezvous. The proposed routing scheme allows a UAV to hover with another UAV or to follow another UAV to increase the communication time. To prevent overall network performance degradation and to allow UAVs to return to the waypoints, this idea estimates energy consumption. The ONE simulator of Helsinki University was used to measure performance.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 499–504, 2020.
https://doi.org/10.1007/978-981-13-9341-9_86

This paper is organized as follows. Section 2 describes the routing schemes currently used in DTN, and Sect. 3 describes the proposed routing scheme. In Sect. 4, the performance of the proposed routing scheme is measured and discussed using the simulator, Finally, Sect. 5 describes the conclusion and future work direction.

2 Related Works

In DTN environment, routing protocols are classified into deterministic routing and stochastic routing [2]. However, since this paper is based on stochastic routing, a description of deterministic routing will be omitted.

The stochastic routing assumes an environment where network changes are unknown. Typical stochastic routing is Epidemic and MaxProp. Epidemic infects a new node by forwarding a copy of the packet whenever the packet-carrying node encounters a node that does not have a copy of packet, like the spread of an infectious disease. When the destination node first meets an infected node, it receives the packet [3].

The MaxProp uses several mechanisms to define the order in which packets are transmitted and discarded. In the MaxProp, each node initially sets the probability of meeting all other nodes in the network and updates this value by exchanging it with neighboring nodes. These probability values are used to calculate the path cost to the destination. Each node forwards messages over the lowest cost path. MaxProp also uses a two-part queue which is divided by adaptive threshold. MaxProp sets the message priority with hop count. It assigns a high priority to a new message and forwards it first with low hop count and delete a message with a high path cost if the buffer is full [4].

3 Proposed Routing Scheme

The energy-based routing with rendezvous consists of two modes: ‘request for hovering’ mode and ‘follow’ mode. Both modes are used to increase the delivery probability of important messages, in other words, to increase communication time. Both modes may not be able to transmit all important all messages. Because all UAVs in the network will stop all transmissions and return to the waypoint if only the energy remaining available to return to the waypoint remains. This is also why the opponent UAV does not have to worry about accepting the hovering request. The proposed routing scheme assumes that GPS and path information are exchanged via Hello packets. The message priority is to send the important message first. If the message type is same, the first incoming message is sent first.

Frequent use of modes can cause overall network performance degradation and UAVs must return to the waypoints. Therefore, each mode requires several conditions to be activated.

3.1 Energy Consumption Model

The energy consumption for each mode is as follows.

$$E(h) = T(r) \times (x + h) + E'(h) \tag{1}$$

$$E(f) = T(r) \times (x + f) + E'(f) \tag{2}$$

$E(h)$ and $E(f)$ are energy consumption for 'request for hovering' mode and 'follow' mode respectively. $T(r)$ is the time required to transmit all important messages. x is the energy consumption for transmission. f is the energy consumption for flying. h is the energy consumption for hovering. $E'(h)$ and $E'(f)$ are the amount of energy consumed in returning the waypoint after mode release.

3.2 Condition of Mode Activation

Communication time and time taken for UAV to meet the next UAV for the currently met UAV should be predicted. However, since the situation in which a UAV meets another UAV is different every time, the UAV should maintain the following table (Table 1).

Table 1. Example of table that UAV should maintain for prediction.

UAV ID	$T(e)$	T(n)
UAV_0	0	INFINITE
UAV_1	3.5	15
…		
UAV_n	2.5	7

$T(e)$ is the average communication time, that is, the expected communication time. If the UAV has not met the corresponding UAV, $T(e)$ is zero. $T(n)$ means the average time it takes for the UAV to meet next UAV after it has met corresponding UAV. The Common conditions for activation for both modes are as follow.

$$T(r) > T(E) \tag{3}$$

The conditions for 'request for hovering' mode activation are as follows.

$$E(c) - E(h) < T(n) \times f \tag{4}$$

$$E(c) \leq E(t) \times 0.15 \tag{5}$$

The condition for 'follow' mode activation is as follow.

$$E(c) - E(f) < T(n) \times f \tag{6}$$

$E(c)$ is the current energy of the UAV and $E(t)$ is the total energy of the UAV.

4 Performance Evaluation

This experiment is carried out using ONE simulator. In the experiment, the delivery probability of messages was measured as the size of important messages increased.

4.1 Simulation Results

Network-related parameters such as transmission range, data rate, and average current draw are set based on IEEE 802.11b [5]. The specific simulation environment is shown in Table 2 (Figs. 1 and 2).

Table 2. Simulation parameters.

Parameter	Value
Transmission range (m)	140
Data rate (Mbps)	11
Idle mode (mA)	156
Transmit mode (mA)	284
Receive mode (mA)	190
Buffer size (M)	2,100
Area (m^2)	514,098.3
Energy capacity (mAh)	3830
Speed (km/h)	25
Maximum flight time (min)	25 (no wind)
Maximum hovering time (min)	27 (no wind)
Number of UAVs	5
Important message generation interval (s)	25–35
Normal message generation interval (s)	25–35

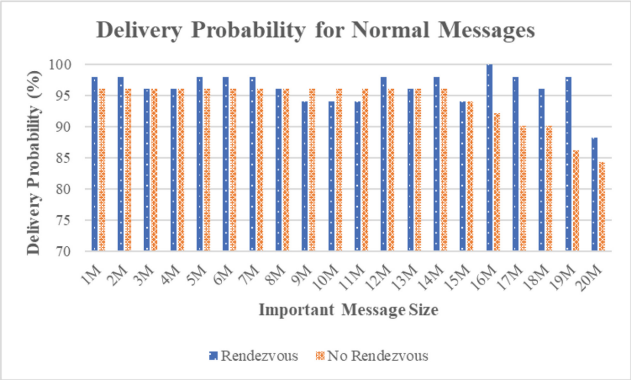


Fig. 1. Comparison of delivery of normal messages with and without energy-based routing using rendezvous.

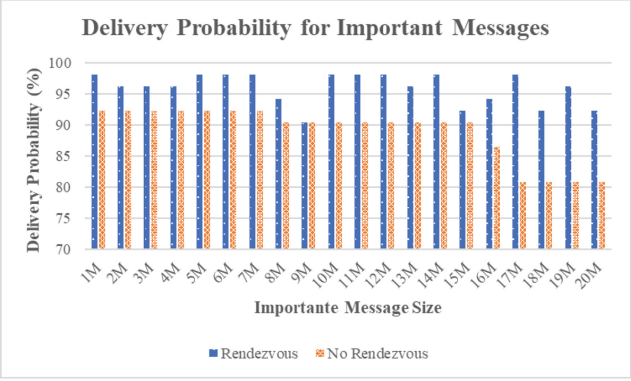


Fig. 2. Comparison of delivery probability of important messages with and without energy-based routing using rendezvous.

4.2 Discussions

Simulation results showed the same delivery probability of important messages when the important message size is 9 M. However, the energy-based routing with rendezvous had a higher delivery probability of important messages in all other cases. When the proposed routing scheme was not applied, the delivery probability of important messages decreased as the size of important messages increased. Especially, it showed that the performance fell sharply from 16 M or more. This means that media data such as high-quality photos are hard to transmit. Also, simulation results showed that the delivery probability of normal messages decreases significantly as the size of important messages increases.

With energy-based routing with rendezvous, the delivery probability of important messages was more than 90.38% even though the size of important messages increased

to 20 M. Therefore, the proposed routing scheme is suitable for multimedia data transmission. In the 'request for hovering' mode, the UAVs hover until all important messages are transmitted, and normal messages are transmitted after the 'request for hovering' mode is released. In the 'follow' mode, the distance between the UAVs may be close to each other depending on the position and movement direction of the UAV, which may result in an increase in the connection time to transmit normal messages after the 'follow' mode is released. Therefore, the delivery probability of normal messages was also good, even though the size of important messages increased.

5 Conclusion and Future Works

Simulation analysis shows that energy-based routing with rendezvous shows robust delivery probability despite the increase in important message size. This feature suggests that the proposed routing scheme is suitable for multimedia transmission. Therefore, it is best to use energy-based routing with rendezvous in scenarios where the size of important messages is large. For example, in the case of a fire monitoring system, it is difficult to confirm the exact fire with the temperature sensor due to high temperature in the summer. In this case, an image file can be more meaningful than hundreds of status messages.

Frequent use of each mode can adversely affect the overall performance of the network. Therefore, the more precisely the situations that need to use the modes are predicted, the better the network performance will be. Future work will be to further refine these predictions to further enhance overall network performance.

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science, ICT & Future Planning) (NO. 2016R1A2B4013118).

References

1. Hyun, S.S., Jeong, H.J., Choi, S.S.: Hybrid spray and wait routing protocol in DTN. J. Internet Comput. Serv. **15**(3), 53–62 (2014)
2. Zhang, Z.: Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges. IEEE Commun. Surv. Tutor. **8**(1), 24–37 (2006)
3. Bindra, H.S., Sangal, A.L.: Performance comparison of RAPID, epidemic and prophet routing protocols for delay tolerant networks. Int. J. Comput. Theory Eng. **4**(2), 314 (2012)
4. Mehto, A., Chawla, M.: Comparing delay tolerant network routing protocols for optimizing L-copies in spray and wait routing for minimum delay. In: Proceedings of the Conference on Advances in Communication and Control Systems, Gwalior, India, pp. 6–8 (2013)
5. Feeney, L.M., Nilsson, M.: Investigating the energy consumption of a wireless network interface in an ad hoc networking environment. In: Proceedings of the IEEE Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2001, vol. 3, pp. 1548–1557. IEEE (2001)

Deep Learning-Based Algorithm for Object Identification in Multimedia

Sangkyun Ko, Bongjae Kim(✉), and Jeong-Dong Kim(✉)

School of Computer Science and Engineering, Sun Moon University, Asan-si, Chungchengnam-do 31460, Korea
highsgl9101@gmail.com, bjkim0422@gmail.com, kjdvhu@gmail.com

Abstract. As the number of multimedia information such as video and images increases in daily life, researches for efficient identification and retrieval of vast multimedia data have shown great results. In this paper, we propose an algorithm based on YOLO framework to accurately identify and retrieve the object that the user wants to find. The proposed deep learning algorithm based on the YOLO framework identifies objects by clearly grasping the intention of users in the multimedia using the web, and provides more accurate image search by supplementing the limitations of the existing CBIR search. These functions can be expected to increase the use of large-capacity video files such as CCTV and automobile black box in recent years, and it is unnecessary to search for a scene where an object desired to be searched by a user is watched for unnecessary time.

Keywords: Deep learning · Object identification · Object detecting · YOLO framework · Object detecting algorithm

1 Introduction

The definition of machine learning refers to the development of algorithms and techniques that enable machines to learn in a similar way to humans. Deep learning is a representative technology for realizing machine learning [1, 2]. Deep learning can be defined as a set of machine learning algorithms using nonlinear transformation as one of the detailed methods of machine learning. Deep learning algorithms include Deep Neural Network (DNN), Convolutional Neural Network (CNN), and Deep Belief Network (DBN) [3]. Deep learning technology has been developed by using deep learning algorithm in computer vision, natural language processing, speech recognition, object classification, motion identification and so on, according to the rapid development of server environment including GPU (Graphics Processing Unit) and the number of patients is increasing [4, 5].

In this paper, we propose a deep learning based algorithm and system for object identification in multimedia. The core technology of deep learning based algorithm and system for object identification in proposed multimedia is real-time object detection (YOLO) which is real-time multi-object classification framework [6]. Using YOLO, various objects can be extracted from multimedia such as photographs and videos. When a user uploads a photograph, the most frequently appearing objects in the

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 505–511, 2020.
https://doi.org/10.1007/978-981-13-9341-9_87

photograph are identified and returned as search results. Also, if you enter the name of the object you want to find in the movie and movie to search for the object in the movie as in the picture, you can automatically identify the object in the movie by searching the position in the movie where the object appeared.

The proposed object identification algorithm is expected to have the effect of using a large amount of video files such as CCTV or car black box, and it is unnecessary to search a scene where an object that a user wants to see and watch for a video is displayed for an unnecessary time.

2 Proposed Model for Object Identification

Figure 1 shows the overall model of object identification proposed in this paper. The video object retrieval system proposed in this paper consists of *Pre-Processing*, *Darknet-YOLO Framework*, *Object Identification Algorithm* and *Result of Identified Objects*

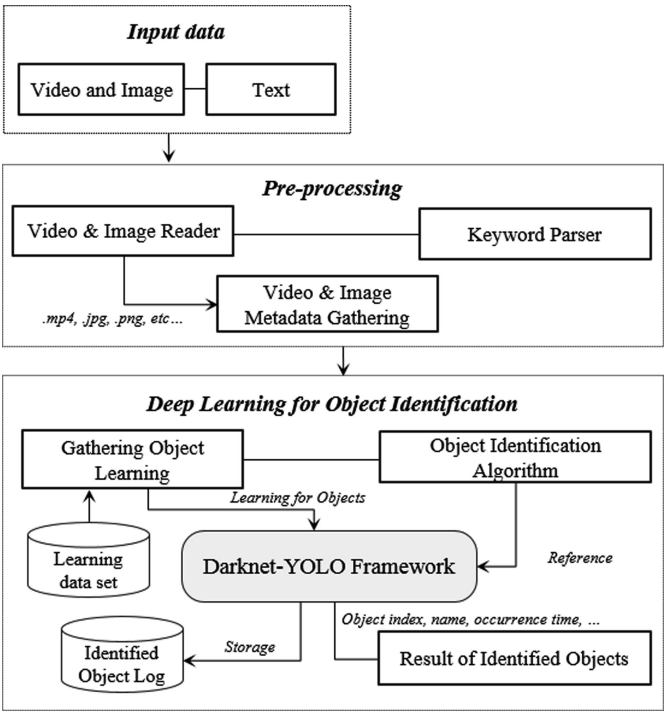


Fig. 1. Conceptual model of processing algorithm for object identification.

Input Data: The data you can upload includes pictures, videos, and text.
Pre-Processing: That are designed for each data will perform a pre-processing step to check whether the uploaded data from the user is in a predetermined data format.
Darknet-YOLO Framework: That is modified and compiled according to the pre-designed intention and learning data set, which is collected data, is collected. When

the image data is received from the user, the object in the image data is extracted and the object classified log file is created by using the image search algorithm.

Object Identification Algorithm: When the image data is received, YOLO formats and counts the extracted object data, and returns the most frequently appearing object in the photograph as the search result. Also, when the video data is received, it analyzes the properties of the uploaded file to obtain FPS (Frames Per Second) and playback time information. After that, the object extracted from YOLO is formatted according to the algorithm and returns the object appearance information to the user as a result.

Result of Identified Objects: As a result of object classification, the data format can be text and image. Text is displayed when you return the position of the object in the movie, and when you point to the object you searched for. The result of the photograph can be returned as a result of a keyword search or an image search for searching for a photograph as a photograph.

3 Object Identification Algorithm

The *Object Identification Algorithm* defines algorithms 1 and 2 for searching objects in images and videos. First, when the user uploads the image, the image is stored on the server and YOLO is executed to extract the object of the uploaded image. Then, in order to know the number of objects appearing in the extracted object information (object), the log file is read and the name of the most frequently appearing object is obtained. If the server has an image matching the most frequently occurring object, it returns the result as a result.

Algorithm 1. Object identification in images

```
OUTPUT
    I search result image
METHOD:
    run D(J)
    open L, read L
    N = count the most appear objects in the picture use L
    M = object name with highest N
    if ID has M then
      return I
    end if
```

J: user input image file, *L*: object log file, *D*: Darknet YOLO .exe file, *ID*: image data, *M*: most visible objects name, *N*: number of objects

Algorithm 2. Object identification in videos

```
OUTPUT
    I index of object in mp4 file
METHOD:
    run D (M)
    open L
        while i == T do
            OL = read i^th line of the L
            if O in OL then
                F = i
                OA = F / S
                I = The O appeared for OA seconds.
                    return I
                else
                    continue
                end if
                    i = i + 1
            end while
        else
    end if
```

O: object name, ***P***: mp4 play time, ***M***: user input mp4 file, ***F***: object appear frame number, ***D***: Darknet YOLO .exe file, ***S***: fps of mp4 file, ***L***: object log file, ***T***: total number of frames, ***OA***: time when the object appeared, ***OL***: line of object,

4 Experimental and Comparative Evaluation

This section describes the experimental environment, dataset, and experimental results on the actual implementation of the proposed system, and also describes the comparative evaluation.

4.1 Experience Environment

Table 1. Experience environment.

Environment		Description
GPU	GTX 1080 ti * 4	Multi-GPU Environment
CPU	i7-6850K Broadwell-E	Processor
RAM	DDR4 PC4-192000 8G * 8	Main Memory
SW	HTML, JS, CSS	UI implementation
	Flask	Web server implementation
	Darknet-YOLOv3	Deep Learning Framework
	openCV 2.4	Image Processing
	cuDNN 5.1	Accelerated Deep Learning
	CUDA 8.0	GPU Programming
	python 2.4	Server Algorithm Implementation
	Ubuntu 16.04	Operating System

Table 1 shows the development environment of algorithm and system based on deep learning for object identification in multimedia proposed in this paper. In the development environment, a multi-GPU environment is constructed to enhance the fps performance of real-time object detection, and a web service is provided using a web framework Flask.

The data sets used for classification of objects were about 2500 pieces of domestic vehicles i40, Santafe, and Ray which were collected directly. The list of classifiable objects is classified into 90 kinds of objects which are classes of directly collected data 'i40', 'Santafe', 'Ray' and coco dataset [7]. The training data of 2500 Korean domestic vehicles took an average of 7 h for the learning time of 45000 GTX 1080 Ti * 4 based on 50,000 times learning. The initial learning process, the 23rd learning, has a total loss rate of 8.57 and an average loss rate of 47.33, which is repeated until the average loss rate is the lowest. The current learning rate defined in the cfg file is 0.000100, and the time required for batch processing is 1.217245 s. The number of learning data used in the 23rd learning is 1472. Information about the learning process is shown in Fig. 2.

```
23: 8.570871, 47.331749 avg, 0.000100 rate, 1.217245 seconds, 1472 images
Loaded: 0.000036 seconds
Region Avg IOU: 0.269321, Class: 0.422137, Obj: 0.002136, No Obj: 0.001434, Avg Recall: 0.125000,  co
unt: 8
Region Avg IOU: 0.391216, Class: 0.468754, Obj: 0.002856, No Obj: 0.001986, Avg Recall: 0.250000,  co
unt: 8
Region Avg IOU: 0.551295, Class: 0.644361, Obj: 0.001239, No Obj: 0.001480, Avg Recall: 0.750000,  co
unt: 8
Region Avg IOU: 0.338240, Class: 0.544331, Obj: 0.001815, No Obj: 0.001450, Avg Recall: 0.125000,  co
unt: 8
Region Avg IOU: 0.327792, Class: 0.438293, Obj: 0.001174, No Obj: 0.001659, Avg Recall: 0.250000,  co
unt: 8
Region Avg IOU: 0.247435, Class: 0.531882, Obj: 0.000671, No Obj: 0.002053, Avg Recall: 0.000000,  co
unt: 8
Region Avg IOU: 0.285794, Class: 0.432118, Obj: 0.000599, No Obj: 0.001622, Avg Recall: 0.125000,  co
unt: 8
Region Avg IOU: 0.253270, Class: 0.468961, Obj: 0.001862, No Obj: 0.001862, Avg Recall: 0.125000,  co
unt: 8
24: 8.707353, 43.469311 avg, 0.000100 rate, 1.251295 seconds, 1536 images
```

Fig. 2. Snapshot of early learning processing.

4.2 Comparative Evaluation

The performance evaluation for the object identification was performed by comparing the video with the frame rate of 30 fps, which is used when shooting in NTSC 1080p or 1080i.

Figure 3 shows the actual playback time and analysis time based on 1280 * 720 video in 1080TI * 4 multi GPU environment. The reason for comparing the playback time and the analysis time is that the analysis time is faster than a person searching for a moving image to reproduce, thereby indicating that an object in the moving image is efficiently retrieved. It is expected that the analysis time will be shorter depending on the environment of the deep learning server.

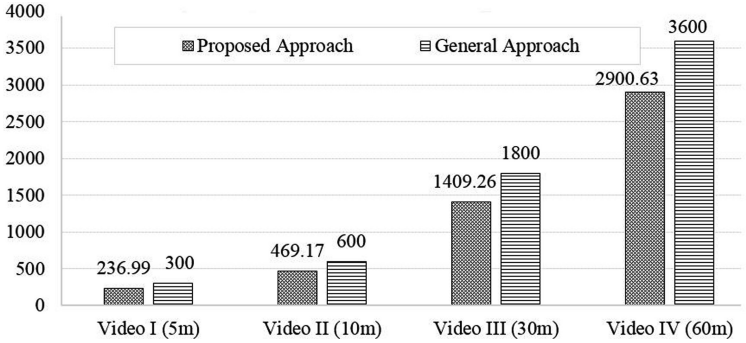


Fig. 3. Comparison of playback time and analysis time for object identification.

5 Conclusion

In order to retrieve objects from multimedia data, it is inconvenient for the user to check one by one and to identify and classify the objects. In order to solve these limitations, we propose an object identification algorithm based on deep learning to support efficient retrieval of object identification in images or moving images. The proposed algorithm uses the YOLO framework to automatically generate a log file for an object in order to extract the object of the image in real time. The proposed object identification algorithm can be used in various fields such as crime prevention using CCTV, fire prevention, traffic control according to the user's purpose. Future researches require additional learning of various objects in order to improve the object recognition rate of the proposed algorithm, and it is required to study the function to search objects in various ways and the reliability improvement of performance.

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2017R1C1B5017476).

References

1. Demirel, B., Cinbis, R.G., Ikizler-Cinbis, N.: Zero-shot object detection by hybrid region embedding (2018). arXiv preprint arXiv:1805.06157
2. Ramachandran, S., George, J., Skaria, S., Varun, V.V.: Using YOLO based deep learning network for real time detection and localization of lung nodules from low dose CT scans. In: International Society for Optics and Photonics, vol. 10575 (2018)
3. GPU. blogs.nvidia.co.kr/2016/02/15/accelerating-ai-artificial-intelligence-gpus/
4. Chung, Y., Ahn, S., Yang, J., Lee, J.: Comparison of deep learning frameworks: about theano, tensorflow, and cognitive toolkit. J. Intell. Inf. Syst. **23**(2), 1–17 (2017)
5. Wu, B., Iandola, F.N., Jin, P.H., Keutzer, K.: SqueezeDet: unified, small, low power fully convolutional neural networks for real-time object detection for autonomous driving. In: CVPR Workshops, pp. 446–454 (2017)

6. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788 (2016)
7. Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Zitnick, C.L.: Microsoft COCO: common objects in context. In: European Conference on Computer Vision, pp. 740–755 (2014)

A Development of Adolescent Depression Screening Using Naïve Bayes Classifier Algorithm

Samuel Garbo[1], Ha-Yeong Kim[1], Syntia Widyayuningtias Putri[1], Ermal Elbasani[1], Hye-Sun Ahn[2], and Jeong-Dong Kim[1(✉)]

[1] School of Computer Science and Engineering, Sun Moon University, Asan-si, Chungchengnam-do 31460, Korea
samuel.garbo@gmail.com, hyk9607@gmail.com, tia.listio@gmail.com, ermal.elbasani@gmail.com, kjdvhu@gmail.com

[2] Graduate School IT Policy and Management, Soongsil University, Sangdo-Ro, Dongjak-Gu, Seoul 06978, Korea
hyesunlas@gmail.com

Abstract. Adolescent depression is a prevalent cause of illness and disability for teenagers. Getting treatment at the earliest sign helps prevent depression from worsening. Depression symptoms can be difficult to tell apart from ups and downs that are part of adolescence thus not being able to contact a professional at the earliest sign. Smartphone applications, are easily available to users of all ages. With an Android application we propose an easily accessible depression screening method for adolescents. Using a widely used depression scale and Naïve Bayes classifier, we compare previous data to the current situation of the user to produce an estimation of the user's situation.

Keywords: Depression · Adolescent depression screening · Naïve Bayes algorithm · Beck depression inventor

1 Introduction

Teen depression is a mental health problem that causes a persistent feeling of sadness and loss of interest in activities. It affects how a teenager thinks, feels and behaves thus causing emotional, functional and physical problems. Teen depression signs and symptoms often include a significant change in thinking and behavior. Depression impairs teens' daily function and if not treated properly it can have long-lasting effects on the person's life. It can be difficult to tell the difference between ups and downs that are part of being a teenager and teen depression and thus may not be recognized at an early stage.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 512–518, 2020.
https://doi.org/10.1007/978-981-13-9341-9_88

Teenagers' anxiety and depression are on the rise. WHO's "Health for the world's adolescents" [1] report reveals that depression is the predominant cause of illness and disability in teenagers aged 10 to 19 years. Some studies show that half of all people who develop mental disorders have their symptoms by the age of 14 [2]. If teenagers with mental health problems, get the care they need, it can prevent death and avoid suffering throughout life.

As smartphones are easily available to people of all ages getting a screening through an application is easily available and is not financially burdensome. In this way it possible to assess depression at an early stage. In this article, we want to propose an adolescent depression screening application which utilizes a widely used depression scale (BDI) and Naïve Bayes Classifier to assess the user's situation.

2 Proposed Model of Adolescent Depression Screening

The screening application consists of a database and smartphone application. The application is developed in the Android platform. The user data and depression data are stored in a database. Figure 1 shows the model of adolescent depress screening.

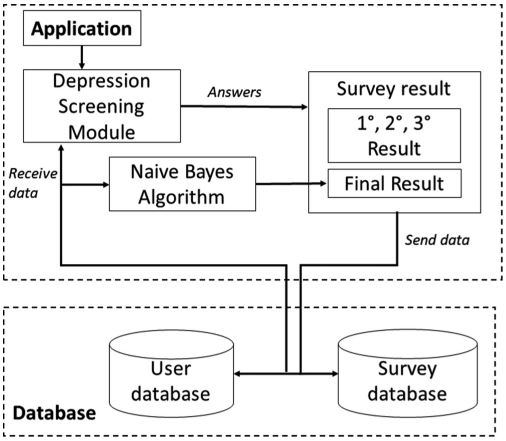


Fig. 1. Conceptual model.

In algorithm 1 the probability for each element is calculated and added to the sum. After calculating the probability of "depressed" and "not depressed" the 2 sums are compared with each other. The algorithm 2 shows an example of calculating the probability of element gender.

Algorithm 1. Naïve Bayes Classification

```
Get data from database
Set sum1 = 0
Set n = number of elements
Repeat n times
        Sum += function of n element [depressed]
End loop
Set sum2 = 0
Repeat n times
        Sum+= function of n element
        [not depressed]
End loop
Compare sum1 and sum2
```

Algorithm 2. Function of Gender for Depressed

```
Function [depressed]:
IF user gender = male
  P (A) = (depressed male + depressed female) / total
  P (B | A) = depressed male / (depressed male + 4: depressed female)
  P (B) = not depressed male + depressed male / total
Else IF user gender = female
  P (A) = (depressed male + depressed female) / total
  P (B | A) = depressed female / (depressed male + depressed female)
  P (B) = not depressed female + depressed female / total
Return P (A) * P (B | A) / P (B)
```

3 Design and Implementation

3.1 Experience Environment and Design

Details for the implementation environment of the application, used for the proposed method of this study are shown in Table 1.

Table 1. Experience environment.

Environment	Description
Android Studio 3.1.3	Android IDE system for application development
Android Virtual Device	Virtual device for application test
Google Firebase Cloud Fire store	Online database for application data storage

Figure 2 shows the main window of the application. The user can start the depression screening by click on the diagnosis menu. After completed all the screening surveys, the user can check the result with result check menu. Also, all the explanation about how to use and screening progress is able to check in the manual menu in the application.

There are 4 steps and 2 types of a survey that the user needs to do before getting the result of the depression screening. This application use Life Event Checklist along with Beck Depression Inventory survey to understand the user's situation.

3.2 Risk Factor Based Screening

According to the literature the cutoff point for depression is different in different age groups, and thus validation studies have been conducted in different countries to determine this specific cutoff point for adult and adolescent populations separately [3]. The first step of this application is a survey to distinguish stressful life events that were assessed through the Life Event Checklist (LEC) with 15 question.

The user will be filling out a checklist question that has 15 questions in the screening form window Fig. 3. Using the answer data and also Naïve Bayes classifier this application will give a result in the last step of the screening progress.

3.3 Beak Depression Inventory Survey

After filling out all the first form of the Risk Factor based screening survey, the user will be directed to the second step which is BDI survey in the second form Fig. 4. Beck Depression Inventory (BDI) is a 21-item, self-rated scale that evaluates key symptoms of depression [4]. The questionnaire was developed from clinical observations of attitudes and symptoms occurring frequently in depressed psychiatric patients and infrequently in non-depressed psychiatric patients [5]. The minimum score is 0 and, the maximum score is 63. Higher scores indicate greater symptom severity. In non-clinical populations, scores above 20 indicate depression [6]. In cases diagnosed with depression, scores of 0–13 indicate minimal depression, 14–19 (mild depression), 20–28 (moderate depression) and 29–63 (severe depression) [7]. Figure 5 shows the result page for the BDI survey. For the second, third and fourth test must be done in one a week. In the third week, the last result page will be coming out and the overall result from the first until the fourth test will be shown on the Result Check page.

Fig. 2. Snapshot of the main menu

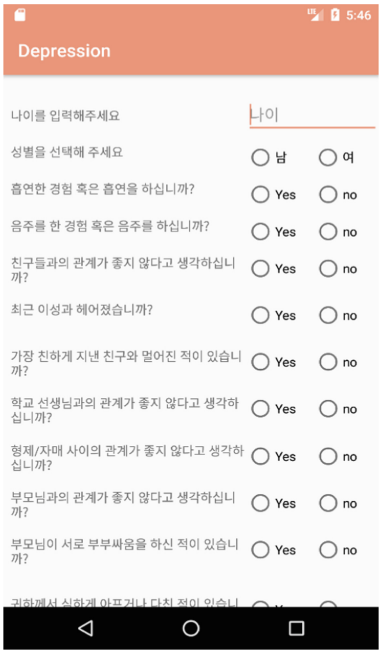


Fig. 3. Snapshot of risk factor checklist

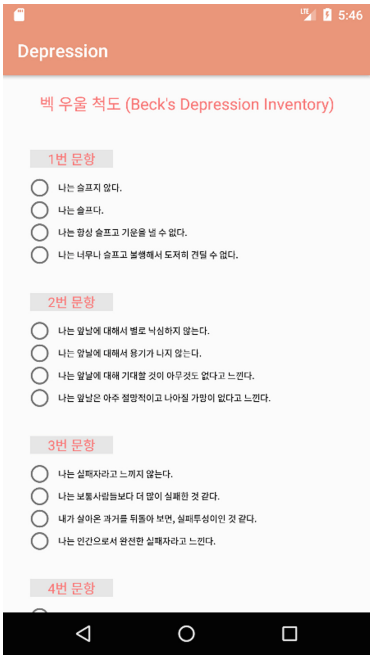


Fig. 4. Snapshot of BDI diagram

Fig. 5. Snapshot of risk factor

3.4 Overall Result

After finishing form the first till the fourth test, all the score and data will be shown in the Result Check page. For the BDI test, the result will be shown in the bar graphs that show the score comparison from the first test until the third test. Also, the result for the Risk factor-based survey will be shown in two categories classified as “depressed” and “not depressed”. In Fig. 6, the bar diagram of the 3 sets of BDI test results, and in the same pages the result of the Risk Factor based survey, displayed as in Fig. 7.

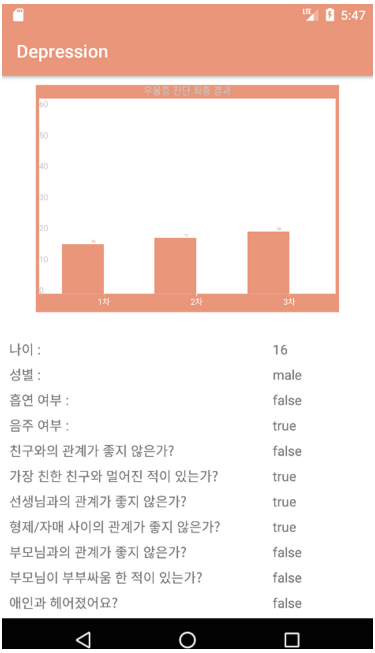


Fig. 6. Snapshot of BDI diagram

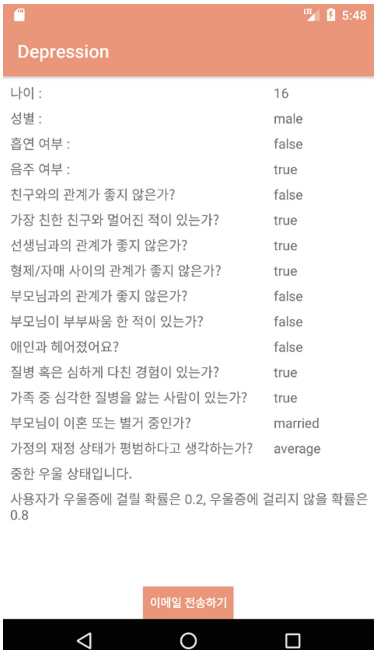


Fig. 7. Snapshot of risk factor

4 Conclusion

The purpose of this application is for early detection of the adolescent depression. With the form of Application that can be downloaded and support a self-screening survey, this application demand is to be helpful for the user for as a reference whether there’s a possibility of depression. The probability measurement based on the real case and survey from many people that diagnosis with depression. This application makes a measurement with the Naive Bayes classifier using the previous data to make a probability of “depressed” and “not depressed” condition.

With two types of survey to give a reliability result that can help the user to raise self-awareness and receive suggestions whether the user needs to get a formal medication or not. Since it is only a self-test and not based on the professional measurement, the user needs to get the further formal diagnosis to from the professional.

Acknowledgments. This work was supported by 2018 LINC+ project of Sun Moon University and by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019R1F1A1058394).

References

1. World Health Organization, Health for the world's adolescent's (2014). http://apps.who.int/adolescent/second-decade/
2. Kessler, R.C., Chiu, W.T., Demler, O., Merikangas, K.R., Walters, E.E.: Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Arch. Gen. Psychiatry **62**(6), 17–27 (2005)
3. Vardanyan, A.: Risk Factor and Prevalence of Adolescent Depression in Yerevan, pp. 30–37. School of Public Health American University of Armenia (2013)
4. Sajatovic, M., Chen, P., Young, R.C.: Rating scales in bipolar disorder. In: Clinical Trial Design Challenges in Mood Disorder, pp. 105–136 (2015). (Chap. 9)
5. Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., Erbaugh, J.: An inventory for measuring depression. Arch. Gen. Psychiatry **4**(561–571), 561–571 (1961)
6. Kendall, P.C., Hollon, S.D., Beck, A.T., Hammen, C.L., Ingram, R.E.: Issues and recommendations regarding use of the Beck Depression Inventory. Cogn. Ther. Res. **11**, 289–299 (1987)
7. Beck, A.T., Steer, R.A.: Manual for the Beck Depression Inventory. The Psychological Corporation, San Antonio (1987)

Survey of Digital Signature Technology for IoT Environment: Focused on KSI's Global Timestamping Technique

Gyeong-Jin Ra and Im-Yeong Lee(✉)

Department of Computer Science Engineering, Soonchunhyang University, Asan-si 31538, Republic of Korea
{rababi, imylee}@sch.ac.kr

Abstract. Digital signatures are security technologies that enable secure communication by providing user authentication, message integrity, and non-repudiation against electronic data in a network. The default behavior of a digital signature is to sign it through the signer's own private key, and everyone can verify it through the public key. Digital signatures are an essential security element in networks and are being studied to suit various environments. With the recent industrial revolution in 4th decade, the advent of a super-connected society like IoT has increased the amount of data communication between networks. Therefore, electronic signature technology using DLT (Distributed Ledger Technology) is being studied as a secure and efficient electronic signature technology for distributed chains such as block chains. In this paper, we describe KSI (Keyless Signature Infrastructure), which is an electronic signature scheme using block chaining as well as various digital signature technologies. Then, the digital signature method is analyzed and compared, and the future direction of the digital signature is described.

Keywords: Block chain · Digital signature · DLT (Distributed Ledger Technology) · KSI (Keyless Signature Infrastructure)

1 Introduction

IoT (Internet of Things) refers to an infrastructure environment that provides convenience to users by communicating with external service providers through devices and the Internet. In this way, the development of IT is developing into a connected society, and the reliability of the whole network is also important through the connection of various objects with secure objects and messages. For this purpose, digital signatures provide user authentication and message integrity and non-repudiation in the communication network. This is provided by three structures and features of digital signatures. First, the digital signature consists of a public-private key. Therefore, the signature is created by using the private key of only the user, and it is provided by the public so that anyone can verify the public key. Second, the public key infrastructure is provided to match the identity of the user with the public key through the public key certificate, and the validity of the public key is provided to guarantee the use of the secure public key. This prevents multiple threats from the network because it is a public

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 519–524, 2020.
https://doi.org/10.1007/978-981-13-9341-9_89

key that anyone can verify. The PKI provides a public key usage environment with a separate trust authority that provides these public key certificates. However, trust organizations must trust the TTP unconditionally in a centralized fashion, and attacks of TTP cause a single point of failure that infringes on user privacy and causes all the harm of downstream users [1]. Therefore, AKI extends the public key usage method to AIP, and public key verification uses a distributed Hash tree to avoid single point error. However, it is not completely distributed and has a disadvantage of adding a surveillance system [2]. Finally, we introduce several types of digital signatures to secure and facilitate the use of digital signatures according to the purpose and the number of users involved. Therefore, various methods are being researched to meet the medium IoT environment according to the features of digital signatures. Among them, public signature technologies such as block chains are being used. In the block chain technology, the same information is always recorded in a distributed form in the form of a signature, which is shared by all of the participants [3]. Therefore, the electronic signature technology using the block chain is a public key-private key that combines the characteristics of user authentication with the public led of the block chain, thereby providing security technology such as digital signature function with the next generation digital signature technology.

Therefore, in this paper, we explain the reason why the electronic signature technology based on block chaining can be used as a suitable digital signature technology in the hyperconnected society like IoT.

2 Blockchain

The block chain is one of the DLT technologies, and data structure technology in which participating nodes both record and carry the same ledger. In other words, the block chain means one transaction list ledger that is distributed. However, unlike existing databases, the ledger can only add transactions and can not be modified or deleted. This is due to two characteristics. First, it has irreversibility due to the hash calculation in the block structure of the block chain. Second, the contents of the previous block form a chain with the contents of the next block, and all the nodes have the same. Therefore, it is very difficult to counterfeit and tamper with reality and has very strong transparency. In addition, since all nodes participating in the block chain store records in the same manner, it is possible to decentralize the network without forming a central organization [4]. Thus avoiding single point of failure in the centralized structure and minimizing broker commissions and procedures. The block chain is divided into a public block chain and a private block chain. A public block chain can be a subject in which all nodes generate blocks and verify each other, assuming that all nodes can not trust each other. Therefore, it has anonymity that information of node is not exposed. On the other hand, private block chains are only the nodes that are verified. Therefore, the block creation speed is fast because only a few specific nodes are required to perform block verification and generation, or to propagate blocks and all nodes have the same general ledger.

3 Digital Signature Infrastructure

3.1 PKI (Public Key Infrastructure)

A public-key-private-key system in which system participants and providers use different keys enables an electronic signature system, including encryption for confidentiality. A PKI is a certificate issued by a trusted issuer and used with a public key to ensure the validity and legitimacy of the public key in a public-key-usage environment. The PKI provides a structure for creating, distributing, using, storing, and revoking a certificate that binds a public key and a proprietary entity [1]. Therefore, client, company and participating agencies etc. in the electronic document storage system sign each private key and encrypt it with the public key of the system. However, a PKI is a centralized system based on TTP that is issued by a trusted personalization agent, and has a SPOF that results in disruption of a few or all services in the event of a failure [2]. In addition, it is highly challenging to detect and control fraud on the part of internal attackers in existing electronic document record systems.

3.2 AKI (Accountable Key Infrastructure)

AKI is a public key verification environment of the Accountable Internet Protocol (AIP), which is a self-authentication protocol. It is supplemented with the PKI's SPOF [2]. It consists of an Integrity Log Server (ILS) and a validator other than the Certificate Authority (CA) of the existing authentication scheme. ILS stores certificates on a hash tree basis; it effectively manages certificates and rapidly addresses data-tampering. Therefore, AKI guarantees the reliability of the certificate, regardless of whether a threat to the secret key occurs through the distributed structure. However, construction of additional ILSs and validators is required, and TLS and additional systems are not fundamentally decentralized with hybrid structures.

3.3 KSI (Keyless Signature Infrastructure)

KSI (Keyless Signature Infrastructure) first appeared in Buldas at el. 2013. It is one of digital signature services through Global Time Stamp provided by Guard Time. The existing PKI - based digital signature is the message and then signs it with the private key and verifies it with the public key [4]. However, KSI uses a hash chain that is repeatedly hashed by SEED as a private key. The public key is used as a pair of the root value of the hash tree in the hash chain and the last value of the hash chain [5, 6]. Afterwards, the chain is used one by one in the opposite direction to the creation, and the user is verified with the feature that only the user knows the value before the hash according to the nature of the hash. The authenticated user sends a KSI server which pairs the hash value of the message with the private key. The KSI server forms a hash tree with the Leaf Node received from a large number of users. According to the hierarchical structure of the KSI, the aggregator collects the information from the client at the bottom, and sends it to the top. The global time stamp and the linked global timestamp are registered by the Publisher [7]. At this time, depending on the type of Publisher, PKI-based TSA token generation issued by an existing trusted authority,

Linked Time Stamping (LTS) using association and association between timestamps, Transient Key using a short trust valid time of modified PKI, Symmetric key) value. Since the initial KSI was published once, it used the LTS method to register it in the media or newspaper so that it was difficult to change it. The KSI-based global timestamp, which uses block-chain technology, provides signature functionality through the notarization of distributed networks. Instead of committing a change to the public media or newspaper of the LTS, the global timestamp is provided as a single block in accordance with the agreement between the publishers [8]. The significance of KSI as an electronic signature is largely the immunity of quantum computing and the global signature of the message. KSI has separate user authentication and message integrity. That is, the integrity of the message is verified by authentication and global hash tree with the key generated by the hash chain [9]. Because all calculations are done in a tree structure of hashes and hashes, the computation is fast and the use of one-time keys is resistive to the theory that existing cryptography is not safe from quantum computers. It also provides signatures for various messages beyond the use of cryptography [10].

(1) Linked Time Stamping based publishing layer KSI
The LTS-based KSI signature constructs a global hash tree from the private key and message pair received from the user, and registers the linking value with the global time to the public media according to the LTS system. The global timestamp value registered in a media such as a newspaper is very strong in terms of stooping/modulating since it is necessary to retrieve and modulate all the media that have been published. In other words, the time stamps issued by LTS are mutually intertwined, and if the integrity of modification and deletion is infringed, the whole structure is changed. Therefore, it is difficult for the attacker to change all the linked stamps to maintain integrity [8]. The structure of the LTS is a linear hash chain and a binary hash tree. The simple linear hash chain-based time stamp scheme hashes the hash value of one document and the hash value of the next document in order as follows, and finally discloses the final hash value in a public media such as a newspaper. An LTS consisting of a binary hash tree is used to construct a relatively large number of documents. Compared to the linear hash tree, the binary hash tree reduces the search time complexity by log, and the overhead is reduced by clustering.

(2) Block Chain based publishing layer KSI
The KSI signature based on the block chain constructs a global hash tree, and commits the global timestamp value linked to the global time to a transaction through consensus among the publishers [9, 10]. The newspaper issued only once a day and the users were not comfortable accessing the media or media. The block chain has a maximum length of 10 min, so everyone has the same ledger, so the availability is improved by improving the accessibility of the user in a relatively short time. Because it is based on block chaining, the security strength and efficiency of KSI depend on the process of agreement and various safety of the block chain. Therefore, it is necessary to improve the structure of the hash tree and the hash strength of the KSI, which is the key constituent of the block chain.

Table 1. Comparison of digital signature infrastructure systems.

Classification	Signature method	Commit method	Key generation method	Signature generation time	Signature verification time	Key usage
PKI	ECDSA	TTP (CA) Public Signature	Discrete logarithm (RSA, ECC)	Slowest	Slowest	Repeat
AKI		TTP (CA) Hash Tree & ILS		Slow	Slow	
KSI	Global Hash Tree & LTS	Open Media, Newspaper etc.	One-way function Hash Chain & Hash Tree	Fastest	Fastest	One-time
	Global Hash Tree & Block Chain	Block Chain		Fastest	Fastest	

4 Analysis

Table 1 compares the PKI signature scheme used in the block chain and the KSI global timestamp as a block-chain application technique. The PKI signature scheme generates an ECDSA signature by generating a key using a lightweight private key-signature key scheme using an elliptic curve. It is classified into Single, Multi, and Threshold signatures depending on the signature scheme. Speed is the fastest Single-Signature, but Threshold-Signature also classifies fragments for a single signature key, resulting in comparable time and efficiency to a single signature. Multi-Signature has the slowest signature generation and verification rate because it has to sign more than n or more than n devices. PKI - based signatures are most commonly used in the block - chain application, cryptography, and generate new account addresses for each transaction to maintain anonymity. However, since the corresponding private key can be used by repeatedly using the same key several times, it focuses on the research direction of generating multiple public account addresses using the same key and safely storing and accessing the private signature key in the hot storage. KSI, on the contrary, is fixed to the root value configured by the hash tree and the private key is used as a disposable one. Therefore, they are stored in Cold Storage and used sequentially one by one. Therefore, considering the future safety of quantum computing, there is a theory that PKI-based public key-private key algorithm is vulnerable. Therefore, KSI has immunity by using one-time key and can perform quick calculation and signature generation using only hash calculation.

Acknowledgments. This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No. 2017-0-00156, The Development of a Secure Framework and Evaluation Method for Blockchain).

References

1. Schukat, M., Cortijo, P.: Public key infrastructures and digital certificates for the internet of things. In: 2015 26th Irish Signals and Systems Conference (ISSC), pp. 1–5. IEEE, June 2015
2. Kim, T.H.J., Huang, L.S., Perrig, A., Jackson, C., Gligor, V.: Accountable key infrastructure (AKI): a proposal for a public-key validation infrastructure. In: Proceedings of the 22nd International Conference on World Wide Web, pp. 679–690. ACM, May 2013
3. Hülsing, A., Gazdag, S.L., Butin, D., Buchmann, J.: Hash-based signatures: an outline for a new standard
4. Nakamoto, S.: Bitcoin: a peer-to-peer electronic cash system (2008)
5. Buldas, A., Laanoja, R., Truu, A.: Efficient quantum-immune keyless signatures with identity. IACR Cryptology ePrint Archive, 2014, 321 (2014)
6. Buldas, A., Laanoja, R., Truu, A.: Efficient implementation of keyless signatures with hash sequence authentication. IACR Cryptology ePrint Archive, 2014, 689 (2014)
7. Emmadi, N., Narumanchi, H.: Reinforcing immutability of permissioned blockchains with keyless signatures' infrastructure. In: Proceedings of the 18th International Conference on Distributed Computing and Networking, p. 46. ACM, January 2017
8. Buldas, A., Kroonmaa, A., Laanoja, R.: Keyless signatures' infrastructure: how to build global distributed hash-trees. In: Nordic Conference on Secure IT Systems, pp. 313–320. Springer, Heidelberg, October 2013
9. Jämthagen, C., Hell, M.: Blockchain-based publishing layer for the keyless signing infrastructure. In: 2016 International IEEE Conferences Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld), pp. 374–381. IEEE, July 2016
10. Buldas, A., Saarepera, M.: Document verification with distributed calendar infrastructure. US Patent 8,719,576 6 May 2014

Decentralized Blockchain-Based Android App Store with P2P File System

Jun-hoo Park(✉), Suh-yu Lee, Geun-young Kim, and Jae-cheol Ryou

Chungnam National University, Daejeon, South Korea
{junhpark, gykim, jcryou}@cnu.ac.kr, seopil7@naver.com

Abstract. In the Android app store, developers must provide a fee of at least 15% of revenue, and there is no reliable evaluation system for apps, so it can lead to occur the damage with installation of malicious file. Blockchain technology which supports decentralized infrastructure can support transactions between developers and buyers without a central organization. Through the Blockchain technology, the fee-free app trading environment can be created. Ethereum's smart contract provides permissions through token issuance and can manages the buyers. Therefore, we implement DAS (Decentralized Android app Store) which provides app trading environment without decentralized organization using smart contract. In addition, IPFS is used for APK file sharing to reduce the amount of gas consumed according to the data size. Since only the hash value of the file is registered, not the whole data of the file, it is possible to consume a little constant average gas costs in the process of file registering and purchasing. We implemented as DApp in Ethereum platform. After purchasing the app, it is possible to induce the appraisal of the app reliability through the evaluation permission assigned to the purchaser, and measuring reliability through the purchase rate and the appraisal rate.

Keywords: Blockchain · App store · IPFS · Smart contract

1 Introduction

An app store is a market where developers can freely upload and users can buy and sell applications. The 'Google Play' is the app store runs in Android environment. It consists of a process in which a developer registers an application, reviewed with 'Google Play' policy and make users to download the app [1]. Google Play is currently the largest app store in the world, with more than 3,600,000 app downloads and $ 20BN of app payments in Q1 2018 [2].

However, in intermediary app store business like 'Google Play', developers must pay more than 15% of their revenue as an intermediary fee for market access [3]. And also, download users in the current app store do not have the privileges more than participant and benefits of downloading app. Furthermore there is no system for evaluating the app, malicious app can be distributed, so a reliable system is required,

Blockchain technology supports a reliable decentralized infrastructure without intermediaries. Smart contract platform such as Ethereum [4] also ensures the reliability of contracts among participants through the developing of decentralized applications

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 525–532, 2020.
https://doi.org/10.1007/978-981-13-9341-9_90

(DApp), In addition, because it is possible to distribute tokens to users who participating in the contract, like Initial Coin Offering (ICO), it can support evaluation authority, compensation system, and so on.

In this paper, we make two main contributions. (1) First, we propose a decentralized Android App Store (DAS) that implements a Android app store which centralized designed, to the Ethereum platform. Based on the smart contract, the developer registers his public key and proceeds the transactions. In order to purchase an app, the user obtains the permission to download by paying the Ether, while at the same time obtaining the token to evaluate the app. This allows app to gain the reliability through the purchase rate and evaluation rate of the app over time.

(2) We use an off-chain file system for efficient sharing of the APK, the Android app file, in a decentralized environment. Ethereum has the capacity of the transaction and gas which is the usage fee of the resource consumed in accordance with the execution of the function but since the gas increased according to the size of the APK file, the off-chain file system IPFS [5] was used. By using IPFS, it is possible to register and distribute APK file with average gas costs regardless of APK file size.

2 Background

2.1 Blockchain and Smart Contract

Ethereum is a crypto currency platform supporting smart contracts which can create an application Dapp (Decentralized Application). Ethereum's Smart Contract supports Turing-completeness and is executed in EVM (Ethereum Virtual Machine) with miner owns [6].

The miner has reached an agreement after executes transaction which received via the P2P network through EVM and participants who trying to execute the Smart Contract pays the gas cost as compensation for the consuming computing resources. The higher the throughput of the transaction and the larger the capacity of the data to be processed, the cost of the gas increases.

Since it is not possible to store all the data in the blockchain, IPFS (Inter Planetary File System) can be used to minimize the gas cost through the off-chain manner to store data in storage and to maintain decentralization. IPFS is the P2P file system and contains DHT (Distributed Hash Table) and Merkle DAG (Directed Acyclic graph) data structure. Because each node saves only necessary file without holding all the files, in the case of file which has large usage, the probability that many nodes have the file is high.

2.2 Android App Signing Structure

"Google Play" utilizes a public key infrastructure to build an application signing structure. When the developer's application is registered, the APK file is registered in the store, and the downloaded user verifies application registered in "Google Play" with its signature.

Currently used basic principle of signing to registered in 'Google Play' is Fig. 1. The pair of public keys used in this process which is divided into two types, application signing keys and uploading keys. The application signing key is used to verify that it is an application registered in "Google Play", and the uploading key is used to prove the personality of the developer.

First, (1) the developer submits the APK file of the application which signed with the uploading private key to Google. (2) At this time, in order to prevent developers from losing their keys, Google stores the developer's application signing private key. (3) Google uses the uploading public key to verify the APK that uploaded by the developer, and (4) signs APK file with stored application signing private key afresh. (5) The download user downloads the APK file and verifies the signature with the application signing public key which contained in APK which generated during signing app procedure. When the verification is completed, the application is installed in the Android environment.

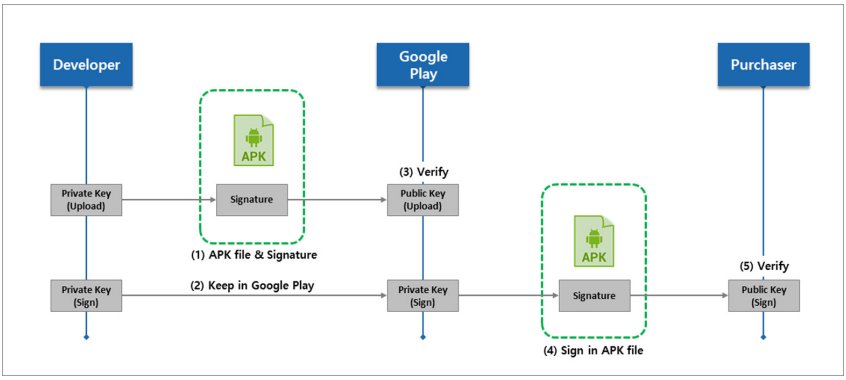


Fig. 1. Andriod app signing structure

3 DAS Architecture

3.1 DAS Entity

In the DAS, roles are assigned to a total of four Entities. A DAS contract that issues a contract corresponding to the public key for each application in place of the role of developer and CA, a Child_DAS contract which corresponding to each application, a Purchaser who download the application. As a Purchaser only, it can be signed for the reliability evaluation of the developer's public key, and it may become Signer when necessary. The content of each Entity is as follows.

Developer: A developer who developed an android application. Set application price, IPFS maintenance cost, Signer distribution cost. When the developer registers the application with Dapp, it invokes the DAS contract in Dapp.

DAS: Contract that plays the role of the top CA. When the developer registers the application, it executes the generateDevSC function and generates the Child_DAS contract based on the value set by developer.

Child_DAS: Contract for each application created by DAS. The public key used for the signature of the application is stored in the contract, and only the Purchaser can evaluate the reliability of the public key. Child_DAS holds the reliability information of the public key of the application.

Purchaser: User who purchases the application. Call the Child_DAS contract, purchase the application, and receive the DHT hash value via Dapp for downloading the APK file from IPFS.

Signer: User who signed the public key of the application. Only purchaser can sign the Child_DAS contract to evaluate the reliability of the application's public key. Purchaser can withdraw his own signature through revocation.

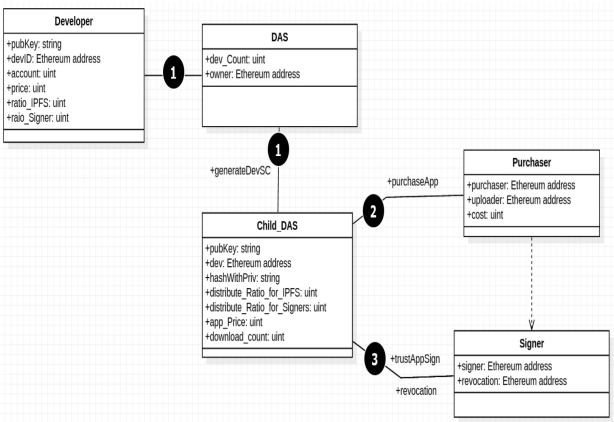


Fig. 2. Program entity UML

DAS is divided into developers who register applications without a central organization, and users who purchase and install applications. In the case of the first user, it is necessary to download directly from the developer node, but as the number of users increases, APK files can be passed from other buyers through IPFS which is a P2P file system. Figure 2. shows the role of each Entity in UML and operates as follows.

❶ Register application and public key: The developer issues transactions to the DAS contract and generates the Child_DAS contract of the application. Child_DAS inherits the DAS contract and operates like an account generated for each application. Register the APK file of the android application in IPFS, and generate a DHT hash value for accessing IPFS.

❷ Payment and application download: (In case of receiving directly from the developer's node) If the user pays Ether to Child_DAS, the Smart Contract

issues a token used for evaluating the application to the purchased user. The token can be used to signing and used to form the trust of the application's public key. The DHT hash value for downloading the file via IPFS is encrypted via Dapp and transmitted.
(In case of receiving from the node, not from the developer) IPFS hash value can be shared because not only developer, but also other purchaser who downloaded knows the value. For example, user C pays to Child DAS Contract to purchase Developer A's application, sets the target to be downloaded to B who is preliminarily purchased, and acquires the APK file from user B.

❸ Signature for reliability: It evaluates the reliability of the application registered by the developer. By signing the public key corresponding to the application by the plurality of buyers, it is possible to provide a decentralized trusted signature system with using the blockchain. It can receive a certain amount of tokens specified by the developer under the condition that maintains the reliability of the application.

3.2 DAS App Signature Structure

The signature system of Android application using smart contract operate like Fig. 3. (1) The developer registers the public key with the remittance of Ether (registration fee) to the DAS contract. (2) The DAS contract generates the Child_DAS contract of the application. In the Child_DAS contract, the public key of the developer is stored in the variable. (3) The developer saves the APK file on IPFS. (4) The user pays Ether to Child_DAS for purchase and registered in the purchaser list. (5) The user who downloaded the APK via IPFS verifies the signature with the developer registered public key in Child_DAS. (6) User who verified download app signs with the public key of the developer (using the token rather than the actual signature) by using the token, thereby the application's reliability increases.

Since the proposed DAS's application signature structure does not have the central authority as the 'Google Play' server, we do not need the public key pair which required for uploading. If a developer uploads an application which containing malicious code, since the user does not vote by using the token, the reliability of the developer does not increase. The lower Child_DAS reliability of the application in proportion to the number of downloads means that it is an unreliable developer, so it is naturally culled in a trusted environment.

4 Implementation

4.1 Smart Contract

Initially when registering the app from DAS, Smart Contract execute generateDevSC() function from DAS to generate Child_DAS. At the time purchaser gives Ether to Child_DAS, purchaseApp() is executed, and a token is generated and assigned to purchaser. Only Purchaser can execute trustAppSign() which can evaluate the

reliability of the app. In addition, a revocation function for withdrawing after evaluation can also be executed. Perform a function to distribute a certain portion of the benefit to purchaser as stored in distribute_for_Signers() to motivate participation in reliability evaluation.

4.2 Application

Developer and user interface for driving DAS Dapp, contract creation for developers, application transactions for purchasers, application reliability evaluation, DHT hash transfer through Whisper [7], application download function via IPFS is provided.

A developer can register an application developed by himself in DAS via Dapp. In this process, generate a transaction that creates Child_DAS contract. The purchaser can buy the registered application, and for the purchased application, the reliability can be displayed. When revenue is generated by a trusted application, at a fixed period, the developer provides reward.

Finally, we will provide the buyer with the DHT hash of the IPFS registered application. At this time DHT hash is provided to purchaser through encrypted channel using Whisper. So purchaser can download the application corresponding to the specified hash with using IPFS via Dapp.

4.3 IPFS

IPFS is used for storing and downloading applications. First, the developer uploads the application to IPFS using 'add' which is a command provided by IPFS. At this time, the DHT hash value of the uploaded application is provided and it is used when searching the APK file in IPFS. After providing a certain Ether, the purchaser receives the DHT hash value of the application through the Dapp from the developer and downloads the APK file from the IPFS using the 'get' command.

Based on 1 gas = 0.000000002 ether [8], 1 ether = 466.74$ [9], in July, Table 1 shows upload gas cost of IPFS applications. When IPFS is not used, when uploading the application directly to Ethereum blockchain, 0.07 $ is added per 100 Bytes of data. Considering that the size of the average Android application is 11 MB [10], uploader need to pay high gas price of 21 $ every time when registering own application.

Table 1. Comparison of gas consumption according to APK file size

Data (Bytes)	Gas	Gas in USD/$	Gas (IPFS)	Gas in USD (IPFS)/$
100	213882	0.19	170325	$0.15
200	281600	0.26	170325	$0.15
300	349504	0.33	170325	$0.15
400	417355	0.40	170325	$0.15
500	485357	0.47	170325	$0.15

Using IPFS, after uploading the application, simply register the corresponding hash value on the Smart Contract. In this case, a hash value (46 Bytes) of the same size is

maintained regardless of the size of the application, so it is possible to provide a hash link of the file with a constant gas charge of 0.15 $. Table 2 shows the amount of gas consumed when an application is registered, purchased, and a function described in Smart Contract such as a reliability signature of the application is called. In the case of uploadApp which registers the application in the block chain, gas consumption of 1.07 $ is required, and in the case of pruchaseApp which is a function to acquire the hash value of the registered application, a gas fee of 0.07 $ is required. A validateApp function that evaluate reliability in the application can also call a Smart Contract through a gas charge of 0.07 $.

Table 2. Gas usage used in smart contract function calls

Operation	Gas	Gas in USD
uploadApp	1163741	$1.08
purchaseApp	83106	$0.07
validateApp	82987	$0.07

5 Conclusion

In this paper, we provided and implemented DAS, an app store that can operate on a decentralized environment. The DAS which developed on the Ethereum Smart Contract shares the APK file with the IPFS, and all interfaces are consisted of DApp.

Since the proposed DAS uses Ethereum crypto currency, by giving a motivation for evaluation by participants of "Google Play" and measuring the reliability from the evaluation, we created system that the application containing the malicious code can be eliminated in stage of trust formation. Our greatest benefits are removing the burden of fees for developers by eliminating central server like "Google Play" and giving the participants the evaluation authority and rewarding the participants. Through our works, we created system that platforms, developers and participants can distribute it's own revenue and take profits.

Acknowledgments. This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (2018-0-00251, Privacy-Preserving and Vulnerability Analysis for Smart Contract).

References

1. Manage app signing keys in Google. https://support.google.com/googleplay/android-developer/answer/7384423?hl=ko
2. Number of available applications in the Google Play Store from December 2009 to September 2018. https://www.statista.com/statistics/266210/number-of-available-applications-in-the-google-play-store/
3. Transaction fees in Google. https://support.google.com/googleplay/android-developer/answer/112622?hl=en

4. Buterin: Ethereum: a next generation smart contract and decentralized application platform. https://github.com/ethereum/wiki/wiki/
5. Off-Chain Data Storage: Ethereum & IPFS, Saving on gas. https://medium.com/@didil/off-chain-data-storage-ethereum-ipfs-570e030432cf
6. Wood, G.: Ethereum: a secure decentralised generalised transaction ledger. Ethereum Project Yellow Paper (2014)
7. Ethereum Whisper. https://github.com/ethereum/go-ethereum/wiki/Whisper
8. Ethereum's information. https://ethstats.net
9. Top 100 Cryptocurrencies by Market Capitalization. https://coinmarketcap.com
10. Average App File Size: Data for Android and iOS Mobile Apps. https://sweetpricing.com/blog/2017/02/average-app-file-size/

A Development of Advanced Communication Platform (ACP) Using Web-AR

Minjeong Song, Sunggeun Yoo, Gooman Park, and Sangil Park(✉)

Department of Information Technology and Media Engineering,
The Graduate School of Nano IT Design Fusion,
Seoul National University of Science and Technology, Seoul, Republic of Korea
mi0_0ng@naver.com, orcogre@gmail.com,
{gmpark, sipark}@seoultech.ac.kr

Abstract. Recently, many AR platforms have been developed by huge global IT companies. However, many AR platforms are not compatible with other platforms and when users get contents, it needed to install heavy software called APPs. Contrarily, Web-AR does capabilities streaming AR contents seamlessly with just Web browsers and it is supporting cross-platform on all devices. In this paper, we have developed a real-time bi-directional communication platform using a marker-based Web-AR called ACP. When the user accesses the ACP and retrieves the 3D object through chat, the matching system generates the QR-Marker. When this QR-Marker is sent to the users, the user can print it or experience the AR through another device.

Keywords: Web-AR · Marker-based AR · Platform · Telepresence

1 Introduction

AR (Augmented Reality) technology refers to the inclusion of virtual elements in views of actual physical environments, in order to create a mixed reality in real time. It supplements and enhances the perceptions humans gain through their senses in the real world [1]. Unlike VR (Virtual Reality), AR has low entry barriers for consumers because they can experience using a smartphone or tablet without a specialized device. For this reason, various global IT companies such as Apple and Google are entering the AR platform market.

However, Most of AR platform is based on applications. APP-based AR requires users to install APPs on each device. Inherently, that solution has not abilities such as cross-platform and automatic AR APP provisioning. Modern software architectures are including concepts like pervasiveness and ubiquitous. Because of that the Web-AR, a promising lightweight and cross-platform approach to AR, is gaining increasing attention owing to its extensive application areas [2].

In this paper, we propose an Advanced Communication Platform (ACP), using Web-AR. By using it, Users who use ACP can access AR contents only by connecting to the web browser without additional software. Also, ACP is a platform using marker-based AR, and its main purpose is telepresence based on real-time user communication unlike existing AR platform.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 533–538, 2020.
https://doi.org/10.1007/978-981-13-9341-9_91

2 Proposed Technology: ACP

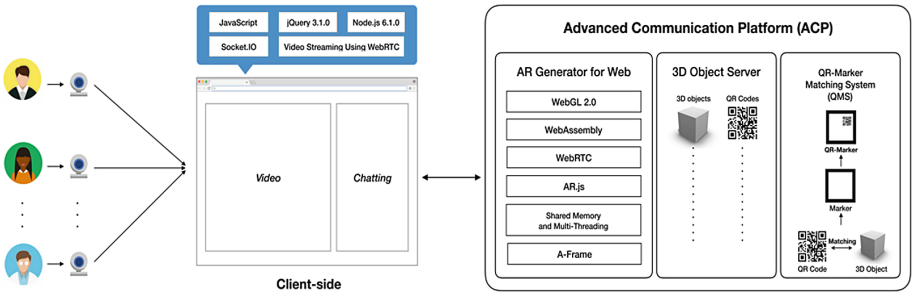


Fig. 1. Configuration of ACP

In this chapter, we discuss the overall system of the ACP. ACP is a communication platform using marker-based AR. As shown in Fig. 1, the system consists of three parts. The first is AR Generator for Web through AR.js developed with WebGL, Three.js, and WebRTC. The second is 3D Object Server stores user-selectable 3D objects and QR codes. The third is QMS (QR-Marker Matching System), it matches QR code and 3D object, generates QR-Marker, and sends it to the client.

2.1 Scenario

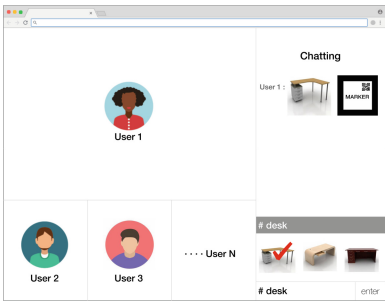


Fig. 2. Client-side of ACP

Figure 2 illustrate the client-side of the ACP. The client-side of ACP has separated to multi-channel communication panel and chatting panel. The Multi-channel panel listed and streamed each user's video screen. The panel applied technologies called WebRTC. It allows each user can access video and audio data through client-side real-time communication with only web browsers. The chatting panel will be posting the user's short messages and images. However, unlike traditional chatting service, It has capabilities searching and posting QR-Markers and 3D objects.

When users are connecting to ACPs, many users can be communicated with other users. Moreover, one of the users searches the chat window using a hashtag, many matched images of 3D objects will be displayed. When the user selected the desired 3D object, the chat window receives the user-selected 3D object and the 1: 1 matched QR-Marker. Each user can display AR on user's screen with QR-Marker.

2.2 Marker-Based AR

Marker-based AR is a technology to show a virtual 3D model by setting white and black 2D image and QR code as visual markers. It is also called Image Recognition or Recognition based AR. In other words, this technology uses the camera in AR device to produce a result. The users get the result when the camera reader senses the marker [3].

2.3 AR.js

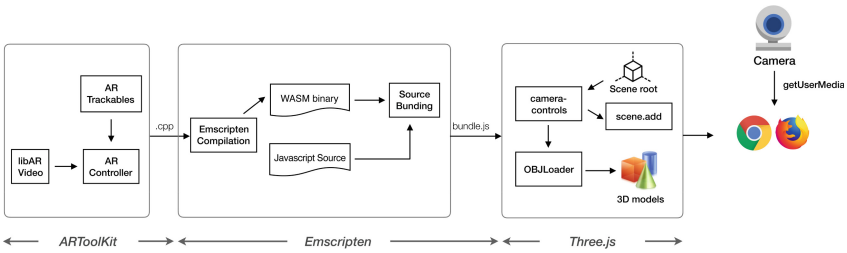


Fig. 3. Configuration of AR.js

AR.js is a solution for efficient processing of augmented reality on the web developed by Jerome Etienne in 2017. It is an open source library developed with three.js and jsartoolkit5-based JavaScript that does not require any other APPs to be installed. AR.js is based on standards and works on any phone with WebGL and WebRTC. It works on Android and Windows mobile [4]. Figure 3 shows the configuration of AR.js.

2.4 QMS (QR-Marker Matching System)

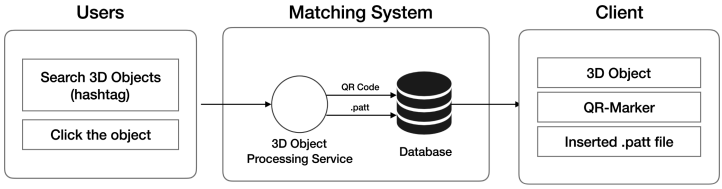


Fig. 4. QMS (QR-Marker Matching System)

As shown in Fig. 4, When a user searches for a desired 3D object using a hashtag and selects a suitable object, it will be transmitted to the Matching System.

The Matching System is composed of a 3D Object Processing Service and database system. A specific 3D object selected by the user goes through the 3D Object Processing Service process and creates a QR code and a marker pattern file (.patt). At this time, the QR code will be mapped 3D objects and markers in a one-to-one relationship. The database system retrieves the QR code and the .patt file to generate the QR-Marker.

The client will be served a 3D object selected by the user, a QR-Marker created through the database, and a .patt file. This system is under patent pending in Republic of Korea.

3 Implementation

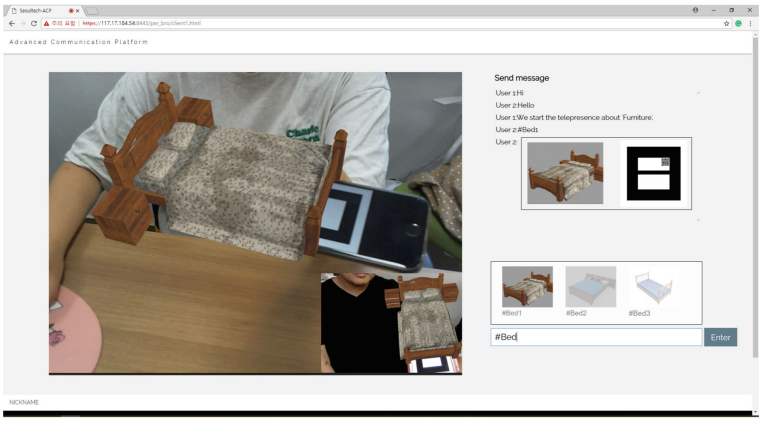


Fig. 5. Implemented ACP

As shown in Fig. 5, both users are connected to the client and communicate with each other. User 2 searches for 'Bed' using a hashtag and displays various 3D objects stored as 'Bed' in the database. The 3D object selected by User 2 and the QR-Marker connected to it are transmitted to the chat window, and User 1 can confirm this. Each user can view the QR-Marker on a webcam via Smart Device and enjoy the AR experience.

The whole client-side system is composed of jQuery 3.1.0 and responsive CSS framework of W3Schools. The real-time chatting data is exchanged with Socket.IO and Node.js 6.1.0 server.

Displaying AR on web browsers is used AR.js library using Mozilla's A-Frame that can work with Three.js. It supports cross-platform, regardless of the device or the user accesses, all with web-based technologies.

The most significant feature of ACP, QMS, was implemented using the MySQL 8.0.11.0 version database. Same as Permalink, QMS creates a unique URI, store the QR code matched with the 3D object, and associate it with the marker to generate the QR-Marker.

With QR-Marker posted in the chat panel via ACP, users can easily access with their smart devices and display ARs. Details of the implementation system will be published as a lengthy paper.

4 Application

Inherently, ACP includes chat and video streaming capabilities, it can be applied to a variety of applications as well as telepresence.

4.1 Application 1: Education

Educational applications have been transformed from a traditional learning into a rich learning experience. For this reason, the educational application is most promising fields when applied advanced technology such as AR. It is known that technological tools used in education provide new opportunities to increase interaction of learners and to learn by enjoying, making the learning process more active, effective, meaningful, and motivating [5].

In historical education, various cultural properties can be shown and explained in detail to students through 3D objects. Concurrently, students can search for and select interesting cultural assets through the ACP and ask questions of the teacher. In addition to history education, chemical molecules, solar systems, etc. can be explained.

4.2 Application 2: Personal Broadcasting

Recently, multimedia services based on social network service, like social TV, personalized multimedia services, and personal broadcasting services are being developed [6]. The most popular part in a personal broadcasting field is gaming. The broadcaster will be explained interactive characters or items in the game with 3D to viewers through ACP. Viewers also have an opportunity to chat with the broadcasters about various 3D objects related to the game and share their opinions.

5 Conclusion

In this paper, we developed a marker-based Web-AR platform ACP capable of real-time bi-directional communication. When a user searches for a desired 3D object using a hashtag through a chatting panel of the client, the QMS existing on the platform matches the selected object with the marker and send it back to the chatting panel. This allows users to easily and conveniently experience the AR without having to install APPs on every device.

It will be applied as a part of education that can enhance students' understanding as well as telepresence. And also it will be providing much more immersive experiences to viewers in personal broadcasts such as YouTube and Twitch.

Furthermore, ACP can be applied to MR (Mixed-Reality) and XR (X-Reality) beyond AR.

Ackowledgements. This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No. 2016-0-00099, Personal Broadcast Technology Development for Production Convenience and Maximum Viewing Experience).

References

1. Di Serio, Á., Ibáñez, M.B., Kloos, C.D.: Impact of an augmented reality system on students' motivation for a visual art course. Comput. Educ. **68**, 586–596 (2013)
2. Qiao, X., Ren, P., Dustdar, S., Chen, J.: A new era for web ar with mobile edge computing. IEEE Internet Comput. **22**(4), 46–55 (2018)
3. What is Augmented Reality? – Types of AR and Future of Augmented Reality. https://dev.to/theninehertz/what-is-augmented-reality–types-of-ar-and-future-of-augmented-reality–1en0
4. AR.js: Efficient Augmented Reality for the Web. https://uploadvr.com/ar-js-efficient-augmented-reality-for-the-web/
5. Yilmaz, R.M.: Augmented Reality Trends in Education Between 2016 and 2017 Years. IntechOpen, London (2018)
6. Seo, D., Kim, S., Jo, Y., Park, H., Ko, H.: Personal multi-angle media broadcasting service system. In: 2012 IEEE International Conference on IEEE Consumer Electronics (ICCE), pp. 715–716 (2012)

Design and Implementation of Real-Time Web Authoring Tool Based on Drag-and-Drop Method

Kichoel Park[1], Boseon Hong[2], Sangkyun Ko[3], and Bongjae Kim[3](✉)

[1] Department of Electronic Engineering, Sun Moon University, 70, Sunmoon-ro 221 Beon-gil, Tangjeong-myeon, Asan-si, Chungcheongnam-do 31460, South Korea
pkcjjal47@gmail.com

[2] Department of Computer and Electronics Convergence Engineering, Sun Moon University, 70, Sunmoon-ro 221 Beon-gil, Tangjeong-myeon, Asan-si, Chungcheongnam-do 31460, South Korea
goodcools34@gmail.com

[3] Division of Computer Science and Engineering, Sun Moon University, 70, Sunmoon-ro 221 Beon-gil, Tangjeong-myeon, Asan-si, Chungcheongnam-do 31460, South Korea
highsgl9l0l@gmail.com, bjkim0422@gmail.com

Abstract. It is not easy for a novice who is not accustomed to coding to create a Web site or to create a Web service. In this process, the process of creating a GUI is a necessary and time-consuming task. Therefore, it is necessary for novice users to have an authoring tool that can easily create and utilize GUIs. In this paper, we design and implement a Web site and Web service authoring tool based on drag and drop method to solve these problems. We applied the MVC (Model-View-Controller) pattern, which is one of the real-time automation models of Web sites. The proposed authoring tool can construct a Web UI for a website or a service by drag-and-drop interaction without any coding work. It also automatically converts the created GUI to HTML code. Therefore, novice users can easily create websites and web services by using the proposed authoring tool.

Keywords: Web authoring tool · Web UI · HTML · Automatic programming

1 Introduction

The Web is a widely used technique to provide various information or services to many users via the Internet [1]. The web has become more popular as mobile devices such as smartphones are widely used. In addition, there are also increasing numbers of people who run their own websites. However, it is not easy for beginners to build their own Web sites or Web services. In addition, configuring the GUI when creating a website is one of the most time-consuming tasks. Web editors are constantly being developed to solve these problems and help users [2, 3]. In addition, various tools are being developed to assist Web authoring [4].

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 539–543, 2020.
https://doi.org/10.1007/978-981-13-9341-9_92

A Web editor is a professional authoring tool on the homepage that allows a user to create an internet homepage [5]. Even a novice user who does not know well the HTML language can easily use it without much coding. In this paper, we propose a drag and drop based real-time web site authoring tool. The main advantage of our authoring tool is that users can easily organize web pages by selecting and placing components for web pages, such as button boxes, radio boxes, and text boxes. It also converts the authored web GUI into HTML code in real-time. Optimization at the code level is also possible using the generated HTML code. Anyway, the proposed authoring tool works by the drag-and-drop method, so it is easy for beginners to use.

The rest of this paper is organized as follows. In Sect. 2, we will explain the design and implementation of the proposed authoring tool. In Sect. 3, we will show the results of the implementation. Finally, we conclude this paper with future works in Sect. 4.

2 Design and Implementation of the Proposed Authoring Tool

In this section, we will describe the design and implementation of the proposed authoring tool. Figure 1 shows the structure of the proposed Web authoring tool. As shown in Fig. 1, the authoring tool proposed in this paper is designed and implemented to operate on Windows operating systems (Window 10, 64 bit) and .NET Framework (version 4.6.1) platform. As we mentioned before, the Model-View-Controller (MVC) pattern, one of the automation programming patterns, has been applied. The proposed Web authoring tool was implemented in C # language.

As shown in Fig. 1, the proposed drag and drop based web authoring tool consists of two main modules. One is drag and drop based UI modeling support module. The other is HTML code auto-generation and real-time display support module.

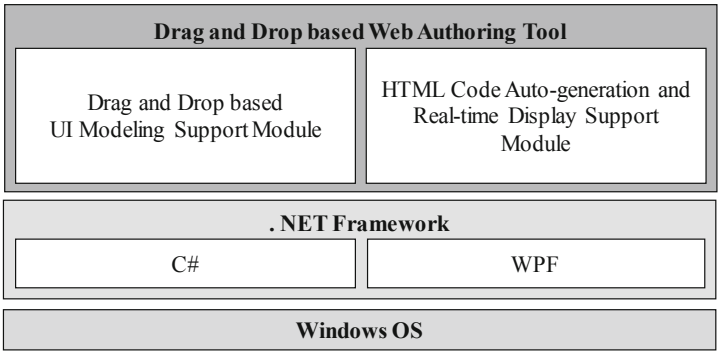


Fig. 1. Structure of the proposed web authoring tool

Figure 2 shows the service flow of the proposed web authoring tool. As shown in Fig. 2, users can make their Web UI based on the drag and drop method. This Web UI is converted to HTML code in real-time by the HTML code auto-generation module. Our web authoring tool also shows the predicted appearance on the web browser in real-time. With this service model, even beginners can easily configure web pages.

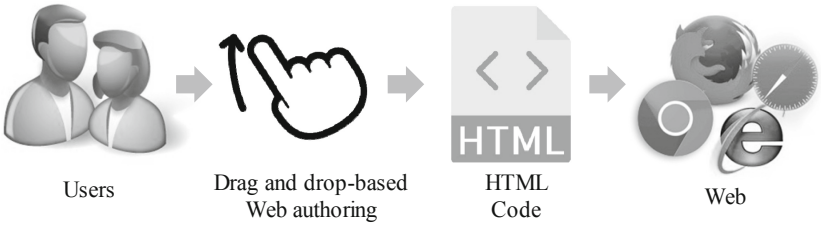


Fig. 2. The service flow of the proposed web authoring tool

3 Implementation Results

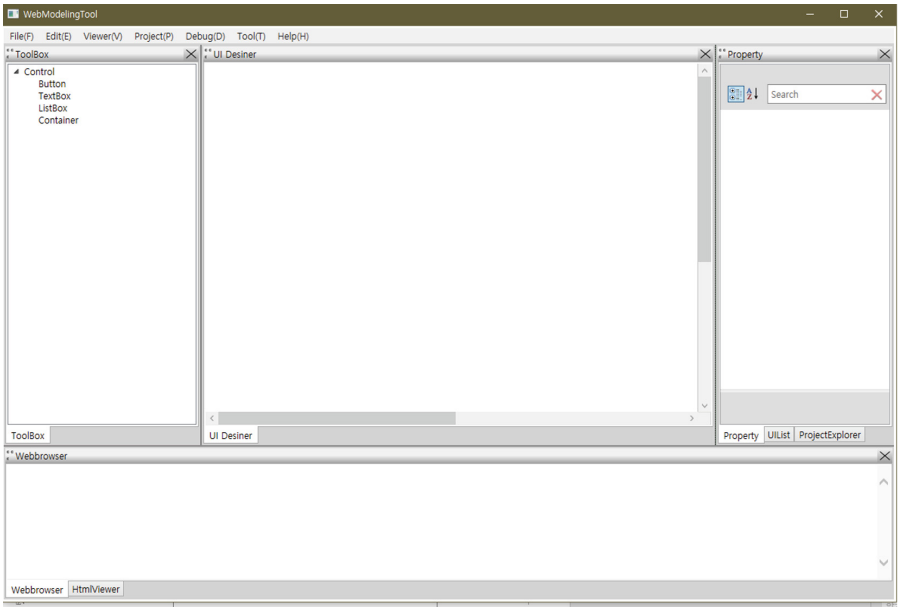


Fig. 3. An initial screen of the implemented web authoring tool

Figure 3 shows the initial screen of the implemented Web authoring tool. The proposed authoring tool is divided into five windows; (1) Top menu window, (2) Tool Box window, (3) Properties window, (4) UI Designer window, and (5) Web Browser window. Top menu window allows the user to select each functionality of the authoring tool. The top menu consists of the six items; File, Edit, Viewer, Project, Debug, Tool, and Help. Tool Box window shows the components that can be used when composing a web page. The current version supports button box, test box, list box, and container, and we will continue to add more GUI components. UI Designer window is a canvas area where drag and drop based authoring takes place. In the Properties window, the user can individually modify the properties such as the name and size of each component of the web page. Web Browser window is divided into two sub-windows. In the HtmlViewer

window, the user can check the HTML code of the authored GUI in real-time. In addition, the user can also see how the authored web GUI will look in a web browser.

Figure 4 shows an example of a Web GUI authoring process based on the drag and drop method. As shown in Fig. 4, there are three Web GUI components; two button boxes and one text box.

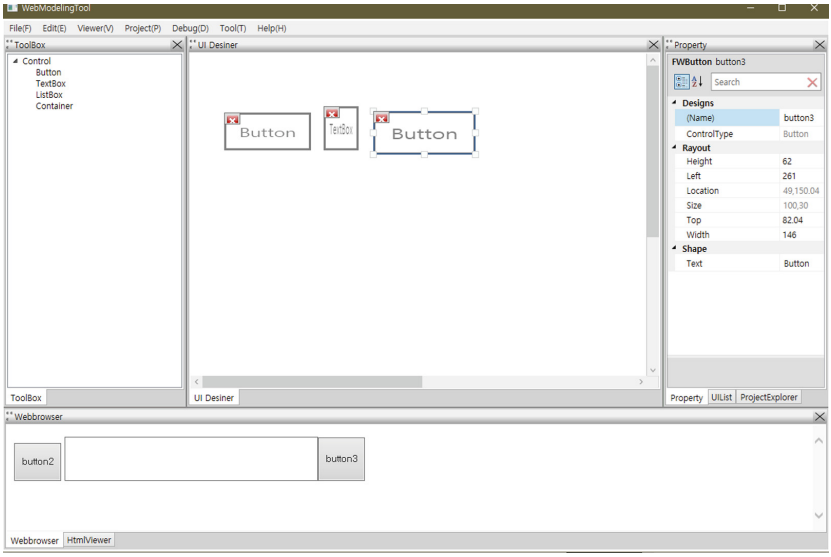


Fig. 4. An example of Web GUI authoring process based on the drag and drop method

Figure 5 shows an example of auto-generated HTML code. The code generated in Fig. 5 is the result obtained through the GUI in Fig. 4. As shown in Fig. 5, we can see that the HTML code is generated correctly. Additional optimizations are possible using the generated HTML code.

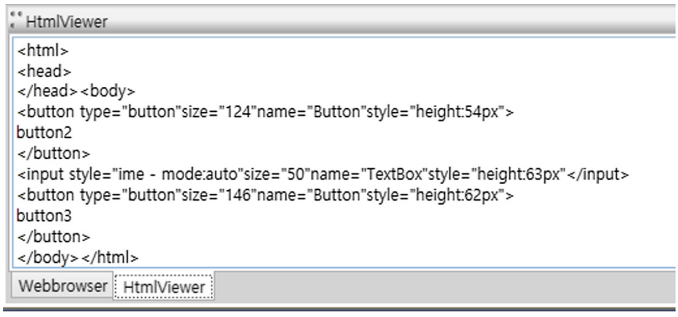


Fig. 5. Auto-generated HTML code verification window

4 Conclusions and Future Works

Web-based services are widely used in various fields because of its accessibility. Web-based services have many advantages. It takes a lot of time to build a website or a Web-based service. There are many repetitive works. Especially, it takes much time to configure the Web UI, and it is difficult for beginners because they are not familiar with HTML code. Various Web editors are being developed to overcome these problems. In this paper, we proposed a real-time Web authoring tool based on drag-and-drop method. Even beginners can easily create the web UI needed to build a website by using the proposed Web authoring tool. Based on the implementation results, we confirmed that the proposed authoring tool can generate the authored Web UI with the corresponding HTML code.

In the future works, we will add more Web UI components into our Web authoring tool to increase versatility and practicality. We will develop a technology that uses deep learning to recognize Web UI sketch images and automatically generate HTML code based on recognition results. The convenience will be increased because Web UI can be created with only a sketch image. In addition, our Web authoring tool will not only be able to make a simple Web UI. We will add the functionality that can model the flow of web services in connection with databases.

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2017R1C1B5017476).

References

1. Nikishkov, G.P., Tsuchimoto, T., Mirenkov, N.N.: Web-based teacher-student interaction in a traditional course. WSEAS Trans. Adv. Eng. Educ. **1**, 15–20 (2004)
2. Zhu, S.S., Liu, W.H., Cai, J.M., Liu, M.W.: The research and design of the web page information system. In: Zhu, H.-S. (ed.) 2015 International Conference on Smart and Sustainable City and Big Data (ICSSC) (2015)
3. Whitt, P.: Website creation software and web browsers. In: Pro Freeware and Open Source Solutions for Business, pp. 169–180 (2015)
4. Chang, K.S.P., Myers, B.A.: WebCrystal: understanding and reusing examples in web authoring. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 3205–3214 (2012)
5. Poley, E.: RUMU editor: a non-WYSIWYG web editor for non-technical users. In: CHI 2010 Extended Abstracts on Human Factors in Computing Systems, pp. 4357–4362 (2010)

An Efficient Method for Wide Area Event Detection and Prediction Using Regression Model

Manh Luong Tien, Ermal Elbasani, and Jae Sung Choi(✉)

Department of Computer Engineering, Sun Moon University, Asan-si, ChungCheongnam-Do 31460, Korea
tien.manh.ptitl994@gmail.com, ermal.elbasani@gmail.com, jschoi@sunmoon.ac.kr

Abstract. Recently, capability of the wireless sensor network is beyond general its purpose with supporting from machine learning technique. Some WSNs deploy in harsh environment where is difficult to recharge energy for sensor, additionally, the spurious event is able to occur that consume a lot of energy to transmit the report message and leading to out of energy soon in WSNs. Therefore energy awareness is the most important consideration aspect of WSNs. In this paper, one method is proposed to reduce energy consumption by recognizing false event detection in cluster head in WSNs and predict burned area of forest adopt regression model.

Keywords: Wireless sensor network · Machine learning · Disaster detection · LEACH-Centralize

1 Introduction

In the last decade, wireless sensor network (WSNs) is an intensive topic which attracts many researchers to investigate about it. WSNs can apply in many areas such as disaster detection, area monitoring, healthcare monitoring and so on. However, disaster detection application is given more consideration rather than others because this application can exploit wide coverage of WSNs. For the reason that it can lead to be tractable in order to detect disaster event in a large area, moreover, there is a huge amount of data coming from WSNs which is feasible to acquire useful information for prediction purpose.

In [1] the author address a system to detect an active volcano that system deploy a wireless sensor network. In these paper, detection system use numerous limited sensor connection via low bandwidth radio, nevertheless, there is no proposal model to recognize when the volcano is active. The event detection depends on number of report from network which is unreliable and inaccuracy. Therefore, to detect event as well as disaster situation a state of art approach is proposed by applying machine learning. Due to the data coming from WSNs is variety and large volume which is very relevant for machine learning algorithm.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 544–549, 2020.
https://doi.org/10.1007/978-981-13-9341-9_93

In this paper, we propose a regression model to predict how much area of forest is burned using data from cluster head in WSNs. The approach separate WSNs into many cluster, each cluster possesses a cluster head which manage its sub-node to refine event reports from sub-node, as well as detecting false event report. Thus, the proposed method solves two main problems in WSNs, the first one is energy consumption and overhead communication between node and base station. Moreover, estimation of burned area also takes into account in this paper by applying machine learning.

2 Proposed System

In WSNs, an event is a phenomenon of the environment as there is something happen, and subsequently affecting to measure data makes it change dramatically. For instance, suddenly, the temperature increase significantly in a short interval, it means there will be a fire happen which makes temperature be higher. Thus, the typical detection way in WSNs is using a threshold but some event is not simple to detect by a threshold, it can be more complicate and depend on several variable. Hence, a distributed approach is applied to recognize an event that is more performance and accuracy.

Overhead communication issue is big problem in the most WSNs which frequently send data from node to base station. To solve this issue, we implement a LEACH-Centralized architecture, this architecture segment the network into N clusters notation as $\{C_1, C_2, C_3, \ldots, C_n\}$ with k nodes. Each cluster head storages the list id address of nodes which is taken into account by cluster head. At the initial time, we suppose the energy of all nodes equal and known node's location which is presented in 2D coordinate as N_i (x_i, y_i), where N_i denotes node number *i-th* and x_i, y_i is its location. And the next assumption is each node known its amount of energy as E_i. A cluster algorithm is performed at base station at initial time of each round, the algorithm is performed at base station, and subsequently base station will send a broadcast message including a list of cluster head to all nodes in network. After received the broadcast message from base station, each node establish a directly communicate with relevant cluster head. Input is all nodes in our system which is divided into k clusters of m nodes in each cluster. The algorithm computes the weight of node based on its information passed into base station, the weight of node is computed following the equation $\omega_i^j = \frac{E_i}{\sum_1^k \sqrt{(x_i - x_n)^2 + (y_i - y_n)^2}}$ which takes into account about the location and energy of individual node inside a cluster. If one node have the weight is maximum, it will be a cluster head. The algorithm returns a list of cluster head that will be sent along with broadcast message.

Event detection plays a vital role in WSNs which decide the amount of reliability of system, as well as accuracy of network. The events are some rapidly changing of a factor of environment which is measured by sensors. In some WSNs, the sensor is deploy for monitoring environment, the environment can be unstable and ambiguous due to climate affection which can influent to measurement data at sensor node. As a result, the false event detection arise that lead to a spurious event detection, this issue need totally considerable as a lot of sensor detect false event, there will be needed a

huge energy to transmit information to base station and the life time of network will be died soon.

As mention above, spurious event need recognize within network as soon as possible be able to reduce transmission energy consumption and extend the life-time of system. To solve the problem as fast as possible, a local coordinator is necessary which take into account all node around it. In the other word, in each local area, there will be a cluster head is picked up to manage sub-node inside its area. The responsibility of cluster head is making decision which event from sub-node is true detection and forwards the data from sub-node to base station. The resource of cluster head is also limited as normal node with a small storage and restrict computational component, thus a light-weight algorithm is performed to refine event report from sub-node, the Distributed Bayesian Algorithm *et al.* [3] is compatible for that scenario.

The Distributed Bayesian Algorithm works on basic idea of probability of neighbor node that is how many nodes having the same measurement data equal evented node in order to make a decision about event. In the real situation, the measurement data can be temperature degree, humidity, wind speed, amount of rain in order to detect fire event. To make it simple, first assumption is we assume all measurement data of node i_{th} is determined as M_i, the variable $M_i = 0$ if sensor detected an event and $M_i = 1$ in else's case. The second assumption is, the variable G_i determines the ground truth at node i_{th} if an event existed $G_i = 0$ and $G_i = 1$ if no event. A false event detection occur when the variable M_i and G_i are diversity as $M_i \neq G_i$ because of either noise or some affected factors. The decision of event recognition in node i_{th} at cluster head is determined as D_i, there are two results of D_i if the cluster head accept event from node i_{th} $D_i = 1$ and reject event $D_i = 0$. Consequently, The probability of decision in cluster head when event is reported from sub-node following the equation:

$$P_a(D_i = 1|M_i = 1, G = 1) = \frac{P(D_i = 1, M_i = 1|G_i = 1)}{P(M_i = 1|G = 1)} \tag{1}$$

$$P_r(D_i = 0|M_i = 1, G = 0) = \frac{P(D_i = 0, M_i = 1|G_i = 0)}{P(M_i = 1|G = 0)} \tag{2}$$

$$P_a + P_r = 1 \tag{3}$$

In forest fire scenario, the sensor nodes are insufficient density, as well as the distance between two contiguous nodes can be several kilometer. Therefore, the measurement data from neighbor node are totally different even the event occurs. To handle this scenario, we assume there are at least k_{min} neighbor nodes that have the same measurement data due to appearing an event at central node. The measurement at k neighbor nodes is computed using equation:

$$P_a(D_i = 1|M_i = 1, G = 1) = \frac{P(D_i = 1, M_i = 1|G_i = 1)}{P(M_i = 1|G = 1)}, \tag{4}$$

where M_{event} is measurement at event node and Δ is the amount of difference between measurement at event node and k neighbor nodes, additionally, we suppose when the

fire event occurs it can influence to neighbor nodes. The impaction is model as spatial correlation of event measurement that means the spatial area around event node there are k nodes which have the identical Δ value in Eq. (5), these nodes is denoted as impact(k, Δ). Hence, the probability of k nodes which have the same value of measurement when an event appears at center node is:

$$P(impact(k, \Delta)|G = 1) = \frac{k}{N}, \tag{5}$$

where k is number nodes have the identical Δ value and N is total node around event node inside a circle with radius = r. Using Theorem 1 in [3], if k is greater than k_{min}, the decision and the measurement at event node will be identical. And Theorem 2 *et al.* [3] if $k \geq 0.5N$ the decision at cluster head is acceptation, and rejection in different cases.

Consequently, when an event is reported to cluster head, it will gather data from N neighbor nodes of event node in order to decide acceptation or rejection report. If the cluster head is acceptable, it will transmit report data to base station using some routing path. At base station the report data is fetched into a regression model to predict an amount of burned area.

In system, input data is variety because of coming from many different kind of sensor, moreover, the requirement of an emergency system alert is precision and real time response. To deal with these problems we proposed a regression model to estimate the amount of burned area for forest fire disaster.

A regression algorithm *et al.* [2] is a linear algorithm works on basic idea of finding a set linear function of coefficient in function $y(x)$ to estimate output y using input x with the smallest loss function E(x), the form of regression model (6) and loss function (7):

$$y(x, w) = w_0 + x_1 w_1 + \ldots + x_D w_D. \tag{6}$$

$$\frac{1}{m}\sum_{i=1}^{m}(h_\theta(x_i) - y_i)^2 \approx E(x). \tag{7}$$

Our system mainly concerns about two tasks such as fire event detection at cluster head and estimate the burned area of forest at base station. To reduce energy consumption, the system automatically chooses cluster head at the initial time of each round depend on the energy-awareness and the location-awareness. When a node is a cluster head, it immediately establishes a one-hop communication with all its sub-node. Whenever a fire event appears at sub-node it will create a report message containing measurement data and transmit to cluster head in order to perform an event recognition task. At cluster head, it has a responsibility to accept or reject an event report from sub-node because some false detection can occur in sub-node due to either noisy or sensor fault. A cluster head is limited resource like sub-node with small storage and computation component, therefore, a distributed detection based on probability is operated cause of the low computation requirement. If cluster head accept a report, it forwards the measurement data to base station using a routing path. In the base station side, when it takes the report data from cluster head, as soon as a regression model is run to predict

how much area burned to broadcast an emergency alarm or evaluate the damage of disaster in real time.

3 Experimental Results

To evaluate the performance between our approaches, we implement a simulation by omnet++ simulator. The scenario consist 99 nodes and one sink, and the nodes cover a square area which is assumed as a surveillance area for forest fire detection. We predefined the number of sub-node in one cluster, each cluster has 8 sub-nodes and one cluster head. Hence, our network is separated into 9 clusters, and in a cluster the distance among sub-nodes is equalization. All nodes have the same amount of initial energy E = 20 mA, furthermore, we make an assumption whenever one node transmits a message, it consume 0.0005 mA/unit distance. Therefore, the energy consumption of node to transmit a message is compute following the equation: $E_{transmit} = 0.0005 * Dis$, where Dis is the distance from source node to destination that uses Euclidean distance. In order to simulation an event at node, a report event message is sent from sub-node to cluster head, at cluster head it will check report event messages. If report event message is "1" which is corresponding to true event detection, the cluster head will forward those messages to sink. On the other hand, if report event message is "0" which mean spurious event detection, the cluster head will reject it. The probability of spurious event detection is 50%.

Figures 1 and 2 show the performance of our method for efficient forest fire detection, the evaluation about two important aspect of WSNs is energy consumption (Fig. 2) and the number of death nodes (Fig. 1). According to those evaluations, the lifetime of WSNs, adopt LEACH-C and event detection, lasts longer approximate 2 times than the lifetime of normal WSNs. Moreover, the number of death node in normal WSNs grow rapidly that can lead to be unreliable in order to detect forest fire event. On the other hand, the number death node of our approach gradually increase, hence the WSNs is more stable.

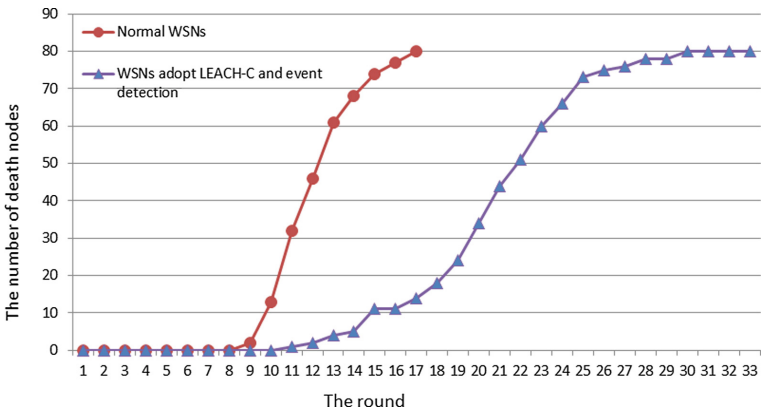


Fig. 1. The total energy consumption each round in simulation

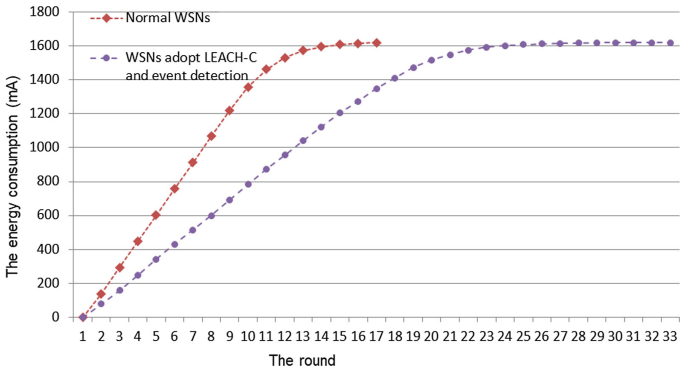


Fig. 2. The number of death node each round in simulation

4 Conclusion

In this paper, we propose an approach to detect false event which can solve overhead communication and energy consumption in WSN. The approach use LEACH-Centralize architecture to separate network into many cluster, and in each cluster implement a probability method to classify event from sub-node. The cluster head spends more energy than other node, hence, after an interval time the system will operate again the LEACH-Centralize to choose the new cluster head. Accordingly, the life time of WSN will be extended as much as cause of energy awareness. Moreover, a prediction model is used in the estimation the forest fire area that can yield the accurate result based on data from WSN.

Acknowledgments. This work is result of studies on the "Leaders in INdustry-university Cooperation" Project, which is supported by the Korean Ministry of Education, and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2018R1C1B5045953).

References

1. Werner-Allen, G., Lorincz, K., Ruiz, M., Marcillo, O., Johnson, J., Lees, J., Welsh, M.: Deploying a wireless sensor network on an active volcano. IEEE internet computing **10**(2), 18–25 (2006)
2. Cortez, P., Morais, A.: A data mining approach to predict forest fires using meteorological data. In: Neves, J., Santos, M.F., Machado, J. (eds.) New Trends in Artificial Intelligence, Proceedings of the 13th EPIA 2007 - Portuguese Conference on Artificial Intelligence, December, Guimaraes, Portugal, pp. 512–523. APPIA (2007). ISBN-13 978-989-95618-0-9
3. Bishop, C.M.: Pattern recognition and machine learning. Springer Science+Business Media, 137–173 (2006)

A Decentralized Consensus Secure and Authentication Framework for Blockchain-Based Healthcare Application

Haider Dhia Zubaydi[1], Yung-Wey Chong[1], Gyu-Sung Ham[3], Kwang-Man Ko[2], and Su-Chong Joo[3](✉)

[1] National Advanced IPv6 Centre, Universiti Sains Malaysia, Penang, Malaysia
haidardhia@yahoo.com, chong@usm.my
[2] Department of Computer Science and Engineering, Sang Ji University, Wonju-si, South Korea
kkman@sangji.ac.kr
[3] Department of Computer Engineering, Won Kwang University, Iksan, South Korea
scjoo@wku.ac.kr

Abstract. The rapid growth of mobile services and Internet of Things (IoT) has caused severe demands of a management system for Mobile Edge Computing (MEC) where User Equipments (UEs) benefit from high computational power, capacity and communication as well as the offered services by MEC. However, a comprehensive management is required to orchestrate the services and resource in MEC to fairly distribute to UEs with the aim of ensuring the Quality of Experience (QoE). In this paper, we propose a decentralization concensus secure and authentication framework for MEC management/orchestration system with crucial security and authentication components by which it ensures the delivery of users' quality of experience.

Keywords: Edge mobile computing · IoT security and authentication · Blockchain · Healthcare

1 Introduction

Blockchain is a platform or architecture that alleviates the reliance on a single or centralized authority. It is designed in a decentralized, distributed manner, that records all transactions across network computers using public digital ledger [1]. Blockchain platform contains a database that stores all transactions using distributed timestamping server and peer-to-peer network, which is autonomously managed. It allows verification and authentication in transactions and user's interactions inexpensively. These transactions are decentralized, immutable, and consensus using cryptography and game theory. Blockchain technology resulted with a robust workflow regarding their data since it provided a high level of security and privacy for participants [2]. The idea of Blockchain was originally proposed by Nakamoto [3] as an electronic cash system, however, it has been merged with new and important margins recently, such as cloud

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 550–556, 2020.
https://doi.org/10.1007/978-981-13-9341-9_94

storage, decentralized markets, authenticated voting, stock exchange, Real estate, healthcare, and Decentralized Apps (DApps) [4].

Bitcoin is the first and most common application of Blockchain technology, which is a currency framework allows two parties to connect directly without any intervention of a centralized intermediary, this process is done via network nodes "miners". Ethereum is a public Blockchain platform that introduced "smart contracts", which are computer code or program that allows two or more parties to control exchange or redistributions of digital assets according to agreements or certain rules that was established between the involved parties [5]. These smart contracts have introduced the concept of decentralized apps (DApps), which provide autonomously operated services stored on the Blockchain for an enhanced interaction procedure.

Healthcare apps are one of the most important domains in current research, it is important to design a healthcare Apps with the ability to connect to technology platforms securely. Because this process includes exchanging patient's data and use the exchanged data across health organizations and app vendors. Today, healthcare researchers suffer from disparate workflow tools, fragmented and soiled data, and delayed communications. On the other hand, healthcare organizations feel reluctant to exchange patient's data due to official regulations and potential financial consequences associated with data sharing [6]. Lack of secure links between independent health systems to establish an end-to-end reachable network is another important key problem. Also, vendor-specific and outdated health systems create gaps in healthcare communications, which makes it harder to improve current systems and introduce a platform to solve all previously mentioned issues [7].

Blockchain platform offers a robust and immune solution to healthcare systems using smart contracts and the software patterns in terms of speed, security, privacy, completeness of fragmented data, context (convert provider-centric systems to patient-centric), trust between providers, and scalability concerns. These solutions can address very serious matters, such as opioid prescription tracking, data sharing between telemedicine and traditional care, patient-controlled cancer data sharing, patient digital identity, personal health records, and health insurance claim adjudication. Using Blockchain in healthcare can really change people's life for the better.

2 Related Works

Applying Blockchain technology on healthcare apps is a recent research area, since the first idea started in 2008, which makes it a very popular domain for creativity and innovation. Many authors proposed different solutions for healthcare systems using Blockchain technology, one way is to create decentralized application leveraging an existing Blockchain such as Hyperledger Fabric [8] by IBM, and Morgan's Quorum [9]. Another approach is done by applying software patterns to address interoperability in Blockchain-based healthcare apps [10]. The authors proposed DASH (DApp for Smart Health) web-based portal for patients that allows them to handle their medical reports. The patient can access and update the medical report, and even submit prescription requests. Some challenges occurred with this implementation when trying to support new users when adding multiple departments within the same organization,

the design was tightly coupled, some resources were duplicated, and the lack of scalability in the designed system. The system requires applying other software patterns to solve other challenges such as security, privacy, dependability, and performance.

A framework design and implementation has been proposed to support electronic medical record sharing for primary patient care [11]. They applied Blockchain technology on oncology-specific clinical data sharing system, this design allows to facilitate the consent management and speed up data transfer using a developed chaincode. It also provides an easily impose fine-gained access control policy for their data to enable efficient data sharing among clinicians. The idea of this prototype has been obtained by MedRec [12, 13], which is a system proposed by MIT, it is based on Ethereum smart contracts for intelligent representations of existing medical records that are stored within individual nodes on the network. It uses Blockchain technology to share healthcare data between Electronic Health Records (EHRs). The issue lies on the fact that researchers are exploring developing the EHR itself using Blockchain as the core infrastructure, which means that all data stored by EHRs will be recorded in a specific block. Which means that these prototypes require huge storage when implemented in massive sized projects.

A recent research proposed an App (called Healthcare Data Gateway (HGD)) [14] architecture to enable patients to easily and securely own control and share their data, without violating official health regulations on patient privacy. The novel privacy risk control model is called purpose-centric access model, which uses unified Indicator-Centric Schema (ICS) to organize healthcare data easily and practically. The issue in this model is the size of stored blocks, since the proposed model store all transactions for record keeping. Keeping all records on Blockchain results with an enormous size chain, which will eventually affect storage network nodes.

Finally, the most recent research has introduced the use of smart contracts as the main method in applying healthcare in Blockchain technology. This system uses private Blockchain based on Ethereum protocol and Internet of Things (IoT) devices [15]. IoT sensors communicates with a device that contains smart contracts to write secure records of different events on the chain. This system supports real-time patient monitoring and medical interventions. This system resolves many security vulnerabilities associated with healthcare systems. Privacy, power consumption, and enormous growth in the number of connected devices concerns arise here since the proposed system uses IoT sensors that require full-time availability of networks coverage.

3 Decentralized Consensus Secure and Authentication Framework for Blockchain-Based Healthcare Application

3.1 Research Overview and Objectives

The main objective of this research is to propose a decentralized consensus secure and authentication framework for blockchain-based healthcare application. The specific objectives are defined as follows: firstly, to propose a decentralized secure framework

with users and IoT device authentication and high-performance remote fog computing capabilities to effectively protect and manage privacy information generated in wearable healthcare IoT environment. Secondly, to propose a storage handling mechanism using software patterns (SecurEdit) in Blockchain platform to manage user's transactions across the proposed framework. Thirdly, to propose a real-time patient monitoring system to secure and maintain the records on framework chains. Lastly, to evaluate the effect of the proposed framework in terms of transaction's speed, security, storage handling ability, and interoperability.

3.2 Wearable Device Virtualization and Authentication

A common device metadata (CDM) for wearable device will be introduced to build data aggregation. The wearable device will sense bio-data from user (i.e. Heart Rate, Spo2, Blood Pressure, Etc.) and these data will be authenticated using smart contract to ensure the security and privacy of user's transactions. The smart contract will connect with the edge framework using SecurEdit when handling the transactions and store them on the database in block manner. In turn, the call will be made to the user's device as alert to handle the process of sharing the information with the hospital if needed (Fig. 1).

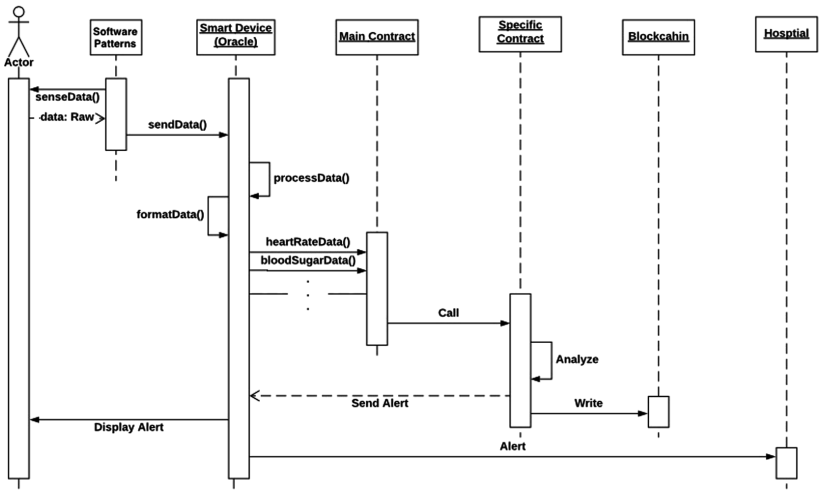


Fig. 1. The logical execution flow of wearable device virtualization [15]

3.3 Blockchain-Based Edge-Centric IoT Computing

At the edge-centric IoT computing, SecurEdit module will be developed. SecurEdit contains the component for monitoring network flow so that personal information will be protected. Distributed agreement technique will be designed to use a cryptographic

token for P2P or D2D information exchange for the healthcare application based on blockchain or Directed Acyclic Graph (DAG). The modules for SecurEdit module is shown in Fig. 2.

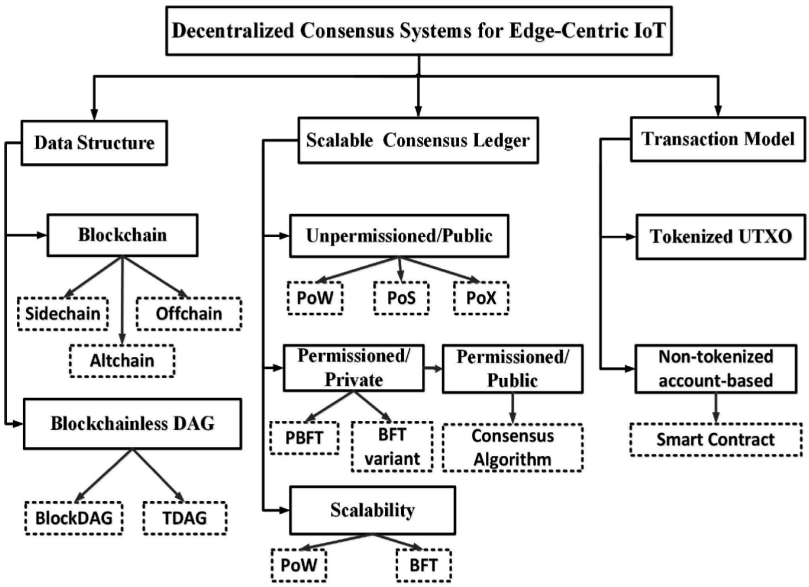


Fig. 2. SecurEdIT core modules: data structure, scalable consensus ledger, and transaction model

3.4 Blockchain-Based Healthcare Application

In order to validate and evaluate the framework we suggested, a cryptoDiabetes application will be develop at the cloud server. CryptoDiabetes will use blockchain technology to protect information about diabetes collected in laboratories and wearable device to prevent counterfeiting and tempering. The application will run on hyperledge together with SecurEdit to validate the effect of a decentralized consensus secure and authentication framework with IoT in terms of transaction's speed, security, storage handling ability, and interoperability (Fig. 3).

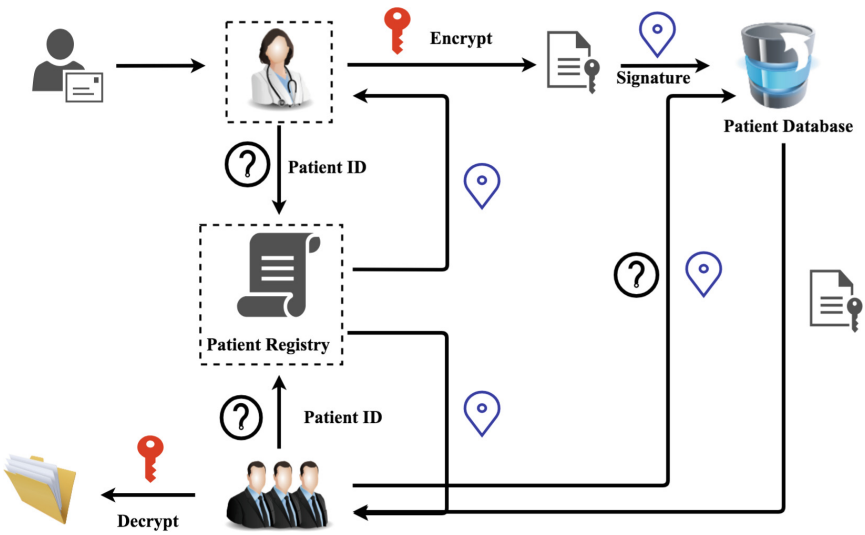


Fig. 3. Blockchain-based healthcare application workflow [10]

4 Conclusion

In this paper, we proposed a new framework to orchestrate the services and resource in MEC to fairly distribute to UEs with the aim of ensuring the QoE for a congested crowd scenario. To design the mentioned framework, we applied two levels of management and coordination such as local for each individual BS and public coordinator for distributed management. In our design we implement resource monitoring/measurement in both levels (Local layer and Public layer) as an essential northbound API over SDN controller to have full visibility of networking resources. We also apply a survey like message sending to UE to store user respond and upgrade the behavior of our framework based on the previous quality of experience.

Acknowledgement. This work was supported by Basic Science Research Program through the Ministry of Education of Republic of Korea and National Research Foundation (NRF) in 2018. (NRF-2017-R1D1A1B03029210).

References

1. Lin, I.C., Liao, T.C.: A survey of blockchain security issues and challenges. IJ Network Secur. **19**(5), 653–659 (2017)
2. Zheng, Z., Xie, S., Dai, H. N., Wang, H.: Blockchain challenges and opportunities: A survey. Work Pap. (2016)
3. Nakamoto, S.: Bitcoin: A peer-to-peer electronic cash system (2008)
4. Johnston, D., Yilmaz, S.O., Kandah, J., Bentenitis, N., Hashemi, F., Gross, R., Mason, S.: The general theory of decentralized applications, dapps. GitHub, 9 June (2014)

5. Buterin, V.: A next-generation smart contract and decentralized application platform. White paper (2014)
6. Downey, A.S., Olson, S. (eds.): Sharing Clinical Research Data: Workshop Summary. National Academies Press, Washington, DC (2013)
7. Adler-Milstein, J., Bates, D.W.: Paperless healthcare: Progress and challenges of an IT-enabled healthcare system. Business Horizons **53**(2), 119–130 (2010). Zhang, P., Schmidt, D.C., White, J., Lenz, G.: Blockchain Technology Use Cases in Healthcare (2010)
8. Hyperledger Fabric (n.d.), 25 July 2018. http://www.hyperledger.org/projects/fabric
9. Quorum (n.d.), 25 July 2018. https://www.jpmorgan.com/global/Quorum
10. Zhang, P., White, J., Schmidt, D.C., Lenz, G.: Applying software patterns to address interoperability in blockchain-based healthcare apps. arXiv preprint arXiv:1706.03700 (2017)
11. Dubovitskaya, A., Xu, Z., Ryu, S., Schumacher, M., Wang, F.: Secure and trustable electronic medical records sharing using blockchain (2017). arXiv:1709.06528
12. Ekblaw, A., Azaria, A., Halamka, J.D., Lippman, A.: A case study for blockchain in healthcare: "medrec" prototype for electronic health records and medical research data. In: Proceedings of IEEE Open & Big Data Conference, vol. 13, p. 13 (2016)
13. Medicalchain, Medicalchain whitepaper 2.1. Technical report Medicalchain (2018)
14. Yue, X., Wang, H., Jin, D., Li, M., Jiang, W.: Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control. J. Med. Syst. **40**(10), 218 (2016)
15. Griggs, K.N., Ossipova, O., Kohlios, C.P., Baccarini, A.N., Howson, E.A., Hayajneh, T.: Healthcare blockchain system using smart contracts for secure automated remote patient monitoring. J. Med. Syst. **42**(7), 130 (2018)

An Educational Data Mining with Bayesian Networks for Analyzing Variables Affecting Parental Attachment

Euihyun Jung(✉)

Department of Convergence Software, Anyang University,
Anyang City, Korea
jung@anyang.ac.kr

Abstract. Despite of its usefulness, Bayesian Networks have been hardly adopted in the educational domain. In this study, we conduct data mining with Bayesian Networks to find which variables are in charge of parental attachment and how much the variables affect. The Hill-Climbing method is selected to learn a Bayesian Network and Markov Blanket is used to extract the related variables to parental attachment. To evaluate the effects of the variables, prediction and diagnostics analyses are performed with the learned Bayesian Network. This paper shows Bayesian Networks are quite effective in finding the related variables and examining the variables' effects in the educational domain.

Keywords: Data mining · Bayesian Network · Educational psychology

1 Introduction

Recently, Bayesian Networks (BNs) [1] have been adopted in various research and industry domains for a wide range of tasks including prediction, anomaly detection, diagnostics, and reasoning [2]. BNs are directed acyclic graph models used to represent knowledge using probability and they allow researchers to easily figure out the relationship among variables. They also enable researchers to observe the trends of a specific variable by changing the probabilities of other variables.

Although BNs have been widely used in many domains to perform data mining and to find meaningful insights, they have been hardly adopted in the educational domain. Educational researchers have preferred statistical methods traditionally and only some of them have just started to use data mining as their research tool in recent. Furthermore, BNs are not the most preferred method even for educational data mining researchers because BNs are relatively new to them and require the considerable knowledge of probability.

Since parental attachment is especially important for the self-identity development and mental health during early adolescence, educational researchers have already conducted a lot of studies [3, 4]. These studies contributed to find related variables and their effects, but the studies had some weaknesses due to their statistical methods. Most of all, they could not handle multivariate data effectively, so they had to make a hypothesis and to select a small set of variables only for evaluating the hypothesis. Since this approach inherently cannot cover all variables, it will miss important

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 557–563, 2020.
https://doi.org/10.1007/978-981-13-9341-9_95

variables which are not considered in the hypothesis. For this reason, many results from the existing educational studies are inconsistent with each other and the studies have a problem to measure the interdependency of multivariate data.

In order to overcome the weaknesses of the conventional approach, we adopt BNs in the educational domain. With BNs, this paper shows how to visually find the related variables to parental attachment and how to measure their interdependencies. In the study, the Hill-Climbing method is selected to learn a BN and Markov Blanket [5] is used to extract the related variables to parental attachment. Also, Netica [6] is used to conduct prediction and diagnostics analyses to examine the effects of explanatory and response variables by changing the variables' probabilities.

The rest of this paper is organized as follows. Section 2 describes the data set in the research and a learned BN model. In Sect. 3, several educational interpretations on the learned BN model are conducted and Sect. 4 concludes the paper.

2 Learning a Bayesian Network Model

2.1 Overview of Dataset

The data for this study was obtained from Korean Youth Panel Survey (KYPS), a longitudinal study of a nationally representative sample of South Korean children and adolescents collected by the National Youth Policy Institute (NYPI) of South Korea [7]. The KYPS was selected using stratified multi-stage clustering and was a nationally representative sample of Korean youth. The KYPS was conducted in five annual waves between 2004 (Wave 1) and 2008 (Wave 5). For this study, the data from the third wave (2006) were analyzed and the children were in the 6th grade classes of an elementary school. Originally, each variable has a 5-point Likert scale (strongly yes, likely yes, neutral, likely no, and strongly no) and the values are mapped to a three-level scale (low, average, and high) for learning a BN model. In this paper, a total of 2,564 subjects were included after dropping out those which had missing fields. The used 33 variables are summarized in Table 1.

Table 1. The description of variables in the dataset.

Description (*Variable name*)
Grade of Language, Mathematics, and English (*gradLEM*), Participation of Language, Mathematics, and English (*partLEM*), Leisure (*leisure*), Frequency of friends' meeting (*meetfrid*), Dad's education (*dad*), Worry about study (*studyworry*), Preference of computer game (*comgame*), Parent's expectation of study (*paexpect*), Interest in school study (*schstudy*), Health (*health*), Spouse abuse (*spouabu*), Parent abuse (*paabu*), Neighbor friendliness (*neiborfrid*), Neighbor's monitoring (*neibormoni*), Self-awareness of bullying (*selfaware*), Surround-awareness of bullying (*suraware*), Worry about surround critics (*surcritic*), Self-esteem (*selfesteem*), Aggressiveness (*agg*), Depression (*depress*), Self-control (*selfcont*), Self-belief (*selfbelief*), Optimistic propensity (*optic*), Stress from study (*stressstudy*), Stress from friend (*stressfrid*), Stress from appearance (*stressapear*), **Stress from material (*stressmaterial*), Parent's monitoring (*pamoni*), Peer attachment (*peerattach*), Stress from parent (*stresspa*), Parent attachment(*paattach*), Life satisfaction (*life*), Teacher attachment (*teaattach*)**

2.2 The Learned Bayesian Network Model

In order to learn a BN model from the dataset, a statistical tool, R is used. We select the Hill-Climbing method as a structural learning algorithm. Since the initial BN model contains too many variables to analyze, we used Markov Blanket to extract the strongly related variables to parental attachment. Using Markov Blanket, we can determine the strongly related variables to parental attachment from the learned BN model as shown in Fig. 1 and the variables are set in boldface in Table 1. This BN model from Markov Blanket shows that parental attachment is influenced by parent monitoring and the stress from parent. Also, it indicates parental attachment affects to life satisfaction, teacher attachment, and peer attachment.

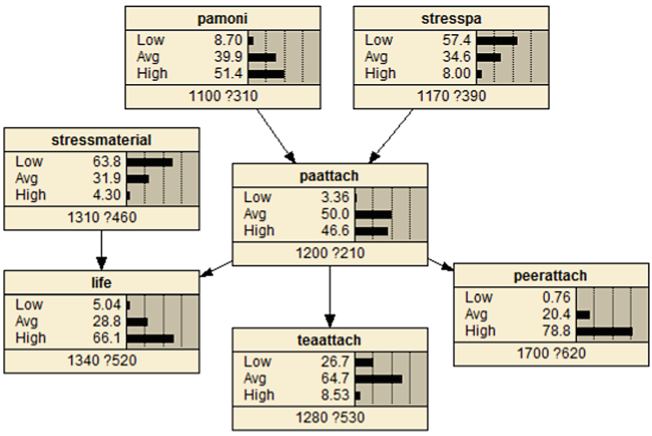


Fig. 1. The BN model from parental attachment's Markov Blanket in Netica

Although Markov Blanket is useful to extract the related variables and to reveal the semantic relations among the variables, it does not explain how much the related variables influence each other. For this purpose, we adopt Netica which shows the conditional probability distribution. Figure 1 shows how the parental attachment's Markov Blanket is displayed in Netica. Each node in the BN has probability values. For example, the "High" value of 8.0 in the *streesspa* node indicates an interviewee may undergoes the stress from parent with the probability of 8.0%.

3 Analysis with the Learned Bayesian Network Model

3.1 Prediction from the Changes of Explanatory Variables

In Netica, it is possible to predict how the response (output) variable is affected by changing the probabilities of explanatory (input) variables. Figure 2 shows the one case of changing the probabilities of the explanatory variables. In the case, we maximize the "High" value of the *pamoni* variable and the "Low" value of the *stresspa* variable.

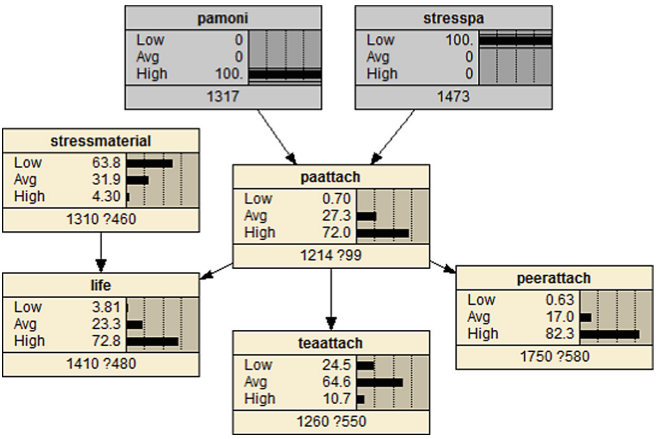


Fig. 2. A case of maximizing the probabilities of the explanatory variables.

In order to observe how much the explanatory variables affect, we conduct experiments by changing the probability values of the explanatory variables. Table 2 shows the nine cases of changing the probabilities of the explanatory variables. When we maximize the "High" value of the *pamoni* variable and the "Low" value of the *stresspa* variable, the "High" value of the *paattach* variable rises up from 46.6% to 72.0% as shown in Fig. 2. This is the highest value of the *paattach* variable from the all cases. It indicates that if parents carefully monitor their child and don't give a stress much to the child, parental attachment can be improved a lot. Reversely, when the explanatory variables have the opposite values, the "Low" value of the *paattach* variable rises from 3.36% to 33.3%. It means that if a child is not monitored and gets severe stress by his/her parents, the child is likely to undergo parental attachment problem with a high probability.

Table 2. All cases of changing the probabilities of the explanatory variables. In the table, "L" means "Low", "A" means "Average", and "H" means "High"

pamoni	*stresspa*	*paatche* (L)	*paatche* (A)	*paatche* (H)
Original	Original	3.36	50.0	46.6
L(100)	L(100)	12.5	64.4	23.1
L(100)	A(100)	11.5	73.6	14.9
L(100)	**H(100)**	**33.3**	**51.5**	**15.2**
A(100)	L(100)	2.6	55.8	41.6
A(100)	A(100)	4.8	72.0	23.2
A(100)	H(100)	13.1	70.2	16.7
H(100)	**L(100)**	**0.7**	**27.3**	**72.0**
H(100)	A(100)	0.5	50.5	49.0
H(100)	H(100)	4.6	54.0	41.4

3.2 Diagnostics on the Given Variables

BNs also enable researchers to conduct diagnostics. When the value of response variable, in other words, an evidence is given, it is possible to determine the likelihood value of the explanatory variables. In Fig. 3, we can assume the case of a child observed to have low peer attachment. Diagnostics with BNs can be useful especially in the educational domain. Frequently, educators cannot observe their students' situations exactly due to practical or political reasons. For example, parental attachment is not easily figured out in school. However, by using diagnostics with BNs, educators can guess parent attachment by examining peer attachment instead.

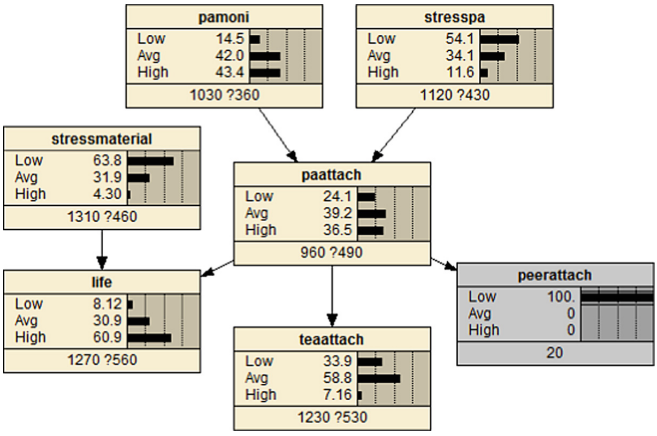


Fig. 3. When an evidence is given, the BN model can determine the probability of the explanatory variables.

We conduct experiments by setting the probability value of the response variables and summarize the results in Table 3. From the experiment, we've found that a child is observed to have low peer attachment, the child is likely to has the problem of parent attachment. As seen in Table 3, the "Low" value of the *paattach* variable rises from 3.36% to 24.1% for this case.

Table 3. This table summarizes the changes of the *paattach's* probability when evidences are given. Each evidence is set to maximize the one of its values; High, Avg., or Low.

Evidence	*paattach (Low)*	*paattach (Average)*	*paattach (High)*
Original	**3.36**	50.0	46.6
life (High: Max)	2.14	41.4	56.4
life (Avg.: Max)	4.49	67.5	27.9
life (Low: Max)	12.9	61.5	22.5
teaattach (High: Max)	0.86	26.9	72.1
teaattach (Avg.: Max)	1.94	52.0	46.0
teaattach (Low: Max)	7.6	52.4	39.9
peerattach (High: Max)	3.04	45.9	51.0
peerattach (Avg.: Max)	3.8	66.0	30.1
peerattach (Low: Max)	**24.1**	39.2	36.5

4 Conclusion

Although BNs have been widely adopted in various domains due to their intuitive graphical representation and support of various data mining, they have been hardly used in educational domain so far. However, if BNs are adopted in the educational domain, they are expected to provide useful insights in the context of education.

Parental attachment plays an important role for self-identity development and mental health during early adolescence, therefore many educational researchers have tried to figure out which variables are important and how much the variables affect with each other. However, with the traditional statistical methods, educational researchers have had a trouble in giving satisfactory explanations especially in the interdependency of multivariate data.

In order to tackle this weakness, we conduct data mining with Bayesian Networks to find which variables are in charge of parental attachment and how much the variables influence. We use Markov Blanket to extract the related variables to parental attachment. The extracted ones are "Stress from material", "Parent's monitoring", "Peer attachment", "Stress from parent", "Life satisfaction", and "Teacher attachment".

To evaluate the effects of the variables, prediction and diagnostics analyses are performed with the learned BN model. From the analyses, we conclude that parents should pay attention to their child and try not to give much stress for maintaining good parental attachment. Also, this study shows that parental attachment can be guessed from the evidence value of its dependent variables such as peer attachment.

References

1. Holmes, D.E., Jain, L.C.: Introduction to Bayesian networks. In: Innovations in Bayesian Networks, pp. 1–5. Springer, Heidelberg (2008)
2. Gandy, A., Veraart, L.A.: A Bayesian methodology for systemic risk assessment in financial networks. Manage. Sci. **63**(12), 4428–4446 (2016)

3. Ma, C.Q., Huebner, E.S.: Attachment relationships and adolescents' life satisfaction: Some relationships matter more to girls than boys. Psychol. Schools **45**(2), 177–190 (2008)
4. Paterson, J., Pryor, J., Field, J.: Adolescent attachment to parents and friends in relation to aspects of self-esteem. J. Youth Adolesc. **24**(3), 365–376 (1995)
5. Tsamardinos, I., Aliferis, C.F., Statnikov, A.R., Statnikov, E.: Algorithms for large scale markov blanket discovery. In: FLAIRS conference, vol. 2, pp. 376–380, May 2003
6. Norsys Software Corporation. Netica. https://www.norsys.com/netica.html
7. National Youth Policy Institute. The 2011 Korea youth panel survey, Seoul, Korea (2011)

LSTM-Based Consumption Type Prediction Model

Jinah Kim[1] and Nammee Moon[2(✉)]

[1] Department of Computer Engineering, Hoseo University, Asan-si, South Korea
jina9406@gmail.com
[2] Division of Computer and Information Engineering, Hoseo University, 20 Hoseo-ro 79 beon-gil, Baebang-eup, Asan-si, South Korea
mnm@hoseo.edu

Abstract. In order to predict consumer behaviors, this paper proposes a consumption type prediction model using LSTM (Long Short Term Memory), a modification algorithm of RNN (Recurrent Neural Network). To do this, we derive and define age and gender consumption type patterns through Association Rule Analysis based on PrefixSpan algorithm using actual card consumption statistical data. Based on this, we used a pattern of daily consumption pattern as an input value, and constructed a model that predicts age and gender consumption patterns by learning the differences between the actual and forecast error rates in a week.

Keywords: LSTM (Long Short Term Memory) · Deep learning · PrefixSpan · Association Rule · Sequential pattern mining

1 Introduction

Predicting consumer behavior is one of the important business strategies for profitability and various studies are currently under way based on machine learning techniques such as CNN (Convolutional Neural Network), XGBoost (Extreme Gradient Boosting), RF (Random Forest), SVM (Supprot Vector Machine) and etc. [1–3]. RNN (Recurrent Neural Network) based research has been carried out since the past because it can not ignore the effect of human consumption behavior on past information [1, 4].

At this time, there is a limitation that there is a long-term dependency problem which is difficult to predict the information of the distant past. LSTM is an RNN-based algorithm for improving it. Recently, research using LSTM as an alternative to RNN has been underway [5–7].

Thus, we propose a consumption type prediction model based on LSTM in this paper. It was constructed using actual card consumption statistical data. We try to derive the pattern of consumption type through association analysis based on the PrefixSpan algorithm, and to predict the consumption pattern by learning the pattern of consumption pattern per day.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 564–567, 2020.
https://doi.org/10.1007/978-981-13-9341-9_97

2 System Architecture

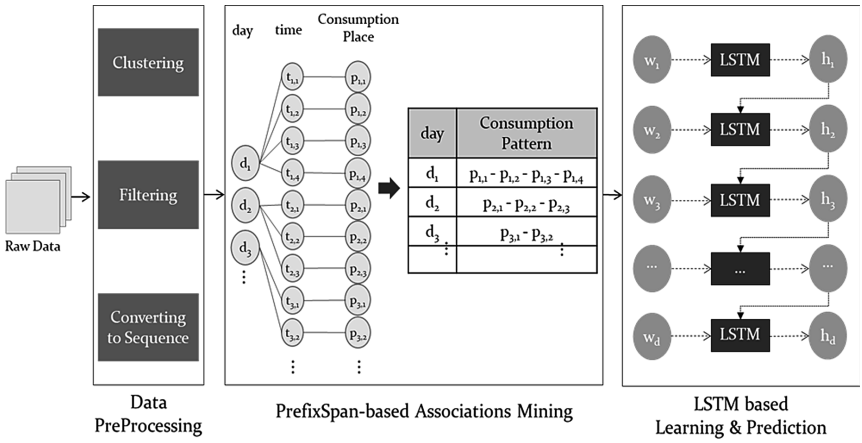


Fig. 1. System architecture

The structure of the proposed system is shown in Fig. 1. First, a data preprocessing process for constructing a model is performed. Consumers are clustered by age and gender by using actual card consumption statistical data, and refinement work is performed for extracting main consumption type. Finally, for association analysis, sort is performed for consumption time and changed into sequence data form.

Sequence type data is used to analyze the association based on the PrefixSpan algorithm for each day. This is a process of finding a frequent item set for a consumption type, which is important for determining the main consumption type. The frequent items found in this process are derived and typified.

Next, LSTM learning and prediction are performed by using association Rule analyzed consumption type. LSTM is performed by predicting the time sequence in a weekly basis, and learning difference between error rate of actual and prediction.

3 The Proposed Method

3.1 PreProcessing

Prior to association analysis, preprocessing involves data clustering, refinement, and transformation. The structure of the initial data consists of age, sex, date, time, industry, and number of consumption. Overlapping occurs at the same time when the population is aggregated in terms of age, sex, date, and time zone. To this end, many are adopted based on the number of consumption, thereby refining the non-consumption type.

The last transformation process is a process of grouping data according to the time sequence with the previously refined data. This is the process of creating a sequence list for future association analysis.

3.2 Consumption Type Association Rule Analysis

Consumption type association analysis is based on PrefixSpan algorithm. Compared with existing Apriori and GSP association analysis algorithms, it is suitable for use in deep learning based models because it searches for frequent patterns without making candidate patterns and improves the problem of memory shortage and time consuming.

This analyzes the pattern of consumption type every day through the sequence data and synthesizes them to define the consumption type pattern. Based on this, the existing data is redefined according to the derived consumption type. The result is used as the input value of the following LSTM based prediction model.

3.3 LSTM-Based Consumption Type Prediction

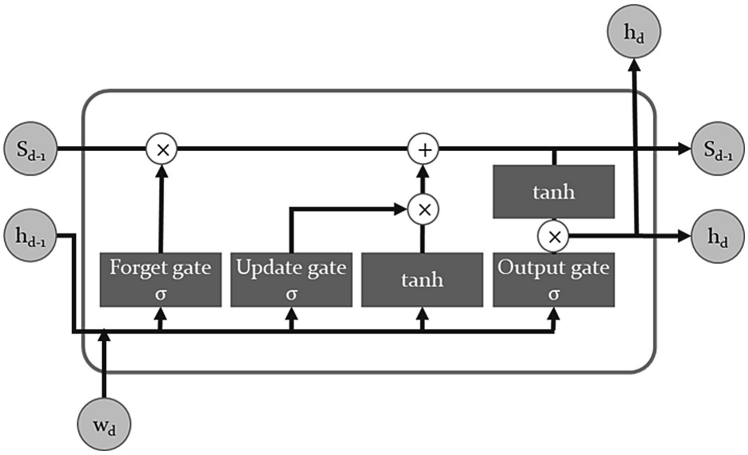


Fig. 2. LSTM cell structure

Based on the results of the above analysis, LSTM-based consumption type forecasting proceeds. The major feature of LSTM is to communicate the learning results of past through Cell State. Therefore, Fig. 2. The computation process first takes a sigmoid for the previous value (h_{d-1}) and the input value (X_d), which is a forget-gate, and determines whether or not to store information between 0 and 1. Next, the value of another sigmoid and tanh function is used as a gate to determine whether the current information is stored in the input gate, and the current information is updated according to whether or not it is stored in the result of the previous step. Finally, the sigmoid takes a value and the tanh function determines which output value to output.

Since the consumption patterns are very different from day to day, the consumption pattern of the weekly unit from Monday to Sunday is processed for the total age (10 to 50) and sex (male, female) Becomes a node having a total of 70 elements. We predict the consumption type after one week and proceed with learning by calculating error rate with actual.

4 Conclusion

In this paper, we propose a consumption - type prediction model based on LSTM for predicting consumers' consumption patterns. LSTM can improve the long-term dependency problem of RNN and is an algorithm suitable for predicting the consumption type over time. In order to obtain the daily consumption type, the association analysis based on the PrefixSpan was performed and used as the input value of the LSTM. We predicted the consumption type after one week by putting it into the input value every week and learned the error rate with the actually analyzed consumption type. In the future, we intend to conduct an expanded study in such a way as to reflect the situation such as climate, consumption trends (SNS, news, media, etc.) that affect consumption patterns.

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government MSIP) (No. 2018020767).

References

1. Krebs, F., Lubascher, B., Moers, T., Schaap, P., Spanakis, G.: Social Emotion Mining Techniques for Facebook Posts Reaction Prediction, arXiv preprint arXiv:1712.03249 (2017)
2. Liu, G., et al.: Repeat buyer prediction for e-commerce. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM (2016)
3. Kim, K., Lee, J.-H.: Predictive models for customer churn using deep learning and boosted decision trees. J. Korean Inst. Intell. Syst. **28**(1), 7–12 (2018)
4. Sheil, H., Rana, O., Reilly, R.: Predicting purchasing intent: Automatic Feature Learning using Recurrent Neural Networks, arXiv preprint arXiv:1807.08207 (2018)
5. Tax, N., Verenich, I., La Rosa, M., Dumas, M.: Predictive business process monitoring with LSTM neural networks. In: International Conference on Advanced Information Systems Engineering, pp. 477–492. Springer, Cham (2017)
6. Nelson, D.M.Q, Pereira, A.C.M., de Oliveira, R.A.: Stock market's price movement prediction with LSTM neural networks. In: 2017 International Joint Conference on Neural Networks (IJCNN). IEEE (2017)
7. Bang, S.-H., Bae, S.-H., Park, H.-K., Jeon, M.-J., Kim, J.-M., Park, Y.-T.: Approach for learning intention prediction model based on recurrent neural network. J. KIISE **45**(4), 360–369 (2018)

Analytic Hierarchy Process Based Cloud Service Assessment Model Using Service Measure Index

Young-Rok Shin, DongKwan You, and Eui-Nam Huh(✉)

Department of Computer Science and Engineering, Kyung Hee University, Yongin 17104, Republic of Korea
johnhuh@khu.ac.kr

Abstract. Cloud service have strength that can increase the satisfaction of customers' various requirements. In that environment, customers should be able to choose the service that best suits their requirements through quality verification. Though service providers use SLAs for quality standards, it is difficult for customers to compare multiple services themselves. Therefore, in this paper, we propose a cloud ranking decision model. For our proposed model, we choose service assessment indicators from SMI model by CSMIC and apply AHP method to measure service scores. And we also show process through a case study how our proposed model assesses cloud services.

Keywords: Cloud computing · Quality of Service · Service measure index · Analytic Hierarchy Process · Service assessment model

1 Introduction

The cloud computing has received much attention from industry and academia. And it is also predicted that will change the economic outlook of computing beyond the technological advances. Cloud service has service environments that can enhance the satisfaction of users' various requirements. In that environment, Cloud Service Providers (CSPs) deliver their service to user with trust in quality. Conversely, Cloud Service Customers (CSCs) should be able to verify the Quality of Service (QoS) and choose an appropriate service provider themselves. However, it is very difficult to compare various services that have already been defined Service Level Agreements (SLAs) [1] with CSP chosen metrics.

To solve mentioned above problem, CSMIC (Cloud Service Measurement Initiative Consortium) proposed SMI (Service Measurement Index) [2]. As of version 1, SMI consists of 7 main characteristics and 51 sub characteristics that form a hierarchical framework. CSMIC expected that the model will help CSCs when they choose a CSP with customization of service indicators [3]. So, many studies have been proceeding using the SMI model. In this paper, we propose a cloud service assessment model. First, we discuss the SMI model to be used in this paper and previous studies using it. Section 3 describes selection of service metrics for assessment and our proposed cloud service assessment model based on the Analytic Hierarchy Process. And

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 568–573, 2020.
https://doi.org/10.1007/978-981-13-9341-9_98

we describe the process of our proposed model through case studies. Finally, we conclude in Sect. 5.

2 Related Works

As a framework, SMI model is a method to calculate the relationship indicator. In 2011, CSMIC announced this model that includes cloud computing specialized QoS indicators called Key Performance Indicator (KPI). So, the KPIs provide as a method to measure QoS and compare with other services. The top level divides the measurement space into 7 categories that are Accountability, Agility, Assurance, Financial, Performance, Security and Privacy, and Usability. Each category is further refined by 4 or more attributes and totally consists of 51 attributes [2].

The traditional weighted sum method for optimization cannot be applied to hierarchical structure. So, many researchers have been studies using SMI model to define SLA. QoE4CLOUD [4] was a proposal of multidimensional framework based on Quality of Experience (QoE). It was proposed Quality centric resource optimization for one or more dimensions that considered about QoS, QoE and business qualities. Another representative research proposed a SaaS quality model [5] that chose and defined some of IaaS quality indicators. However, it was only focused on the QoE based Quality Assessment in [4]. And [5] also did not include the result of experimental evaluation.

Saaty developed an effective Multi-Criteria Decision Making (MCDM) method that Analytic Hierarchy Process (AHP) that provides alternatives through reasonable evaluation. And systematic analysis and stepwise derivation are provided by AHP with pairwise comparison for service quality assessment. Only the quantitative metrics are considered in existing mathematical methodologies. However, the AHP is good to evaluate when the case of considering qualitative metrics. Furthermore, due to the simple and easy, AHP has been used widely for solving MCDM problems. To determine the ranking of various cloud services, we propose a cloud service assessment model based on AHP [6]. After describing the proposed ranking decision model, we will show the process through the case study.

3 AHP Based Cloud Service Ranking Decision Model

3.1 Phase 1: Metrics Hierarchy Based on Chosen KPIs

Figure 1 shows hierarchy of KPI that is based on the mobile cloud service metrics. The goal of this method is located on the top level and this is for finding the appropriate service of all the mobile cloud services that satisfy the requirements of CSC. Quality metrics are contained in the middle level of hierarchical architecture. The bottom level has the mobile cloud service quality values of all the candidate services.

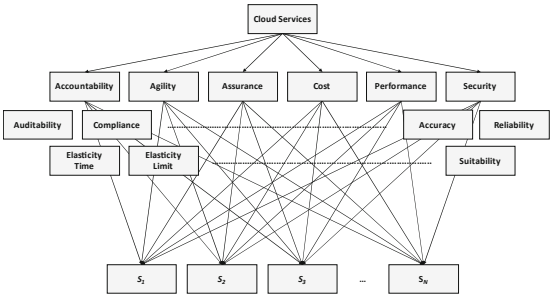


Fig. 1. Metrics hierarchy for mobile cloud service

3.2 Phase 2: Computation of Weights of Each Metrics for Services

To compare various mobile cloud services, we adapted the weight values to each quality value to consider its importance scales. When the CSCs can give the weight values to each metric using numeral values in some scale, for instance 1–9 as suggested in the AHP method [6]. The Table 1 shows the importance of quality metric for pairwise comparison. This methodology was proposed originally for calculating weights for each criterion in the AHP method. This can be used to assign weights to all the QoS attributes. CSCs express their preferences for each metric related to another metrics.

Table 1. Pairwise comparison scale of attributes

Importance degree	Definition
1	Equal importance (no preference)
3	Moderately more importance
5	Strongly importance
7	Very strongly importance
9	Extremely strongly more importance
1/3, 1/5, 1/7, 1/9	Reciprocals of 3, 5, 7 and 9

3.3 Phase 3: Weight Values for Mobile Cloud Service Quality

Phase 3 is to determine weight values for mobile cloud service quality. In our proposed ranking decision model, many metrics were existing, and they made a hierarchy structure by their relationship and its characteristics. For that reason, it is not a simple process to assign the weight values. So, the weight values of each metric need to define using their service performance that is based on the values of the low-level metrics.

To find out the relative importance of different metrics, we construct a pairwise comparison matrix using scale of relative importance. Assuming M metrics, the pairwise comparison of metric i with metric j yields a square matrix $A_{m\times m}$ where r_{ij} denotes the comparative importance of metric i with respect to metric j. In that matrix, $r_{ij} = 1$ when $i = j$ and $r_{ji} = r_{ij}$. Find the relative normalized weight (w_j) of each metric by

(1) calculating the geometric mean of *i*th row and (2) normalizing the geometric means of rows in the comparison matrix. This can be represented as,

$$GM_j = \left\{ \prod_{i=1}^{M} r_{ij} \right\}^{\frac{1}{M}} \quad and \quad w_j = \frac{GM_j}{\sum_{i=1}^{M} GM_j} \tag{1}$$

To geometric mean method of AHP is used in the present work to find out the relative normalized weights of the attributes because of its simplicity and easiness to find out maximum Eigen value and to reduce the inconsistency in judgement.

3.4 Phase 4: Aggregation of Weighted Quality Values for Each Metric

In phase 4, score of each metric needs to aggregate with their own weight values assigned in phase 2. In other words, it is to obtain the overall or composite performance scores for the alternatives by multiplying the relative normalized weight (w_j) of each metric with its corresponding normalized weight value for each alternative and making sum over all the metrics for each alternative.

$$P_i = \sum_{j=1}^{N} w_j \left(n_{ij}\right)_{normal} \tag{2}$$

Where $(n_{ij})_{\text{normal}}$ represents the normalized value of n_{ij}. Pi is the overall or composite score of the alternative A_i. The alternative with the highest value of P_i is considered as the best alternative. This aggregation process is repeated for all the metrics in the hierarchy which results in the ranking of all the cloud services based on KPIs.

4 Case Study of Quality Assessment

In case study, the computation of the service index is done using the QoS data of 3 cloud providers. The QoS data in this case study is collected from other evaluations for 3 IaaS cloud providers [7–9]. We are randomly assigned the unavailable data such as the security level and user required value each service and QoS attribute.

According to SMI model, cloud service has 6 major characteristics that can be matched to mobile cloud characteristics. Finally, we chose the mobile cloud service KPIs that are *Accountability*, *Agility*, *Assurance*, *Cost*, *Performance* and *Security*. As the result of our proposed ranking decision model, each service's final score was 0.2906, 0.2964 and 0.4131 after evaluation. Cloud services ranked as S3 > S2 > S1. However, it is difficult to know the service has cost efficiency or not using only quality values computed as above. CSCs can distinguish only which service has better qualities than the others. Therefore, we will compute the ratio of quality value versus cost. To clearly see which mobile cloud service gives best value at cost, we can visualize it using a graph as Fig. 2. While mobile cloud service S3 has the best overall quality measure, it is also the most expensive. However, mobile cloud service S1 provides the best ratio of quality versus cos (Table 2).

Table 2. Case study example

Top level QoS groups (weights)	First level attributes (weights)	Second level attributes (weights)	Service 1 S1	Service 2 S2	Service 3 S3	User required value
Accountability (0.05)	Level 0–10 (1)	–	4	8	4	4
Agility (0.1)	Capacity (0.6)	CPU (0.5)	9.6	12.8	8.8	6.4 GHz
		Memory (0.3)	15	14	15	10 GB
		Disk (0.2)	1,690	2,040	630	500 GB
	Elasticity (0.4)	Time (1)	80–120	520–780	20–200	60–120 s
Assurance (0.2)	Availability (0.7)	Availability (1)	99.95	99.99	100	99.9%
	Service stability (0.2)	Upload time (0.3)	13.6	15	21	–
		CPU (0.4)	17.9	16	23	–
		Memory (0.3)	7	12	5	–
	Serviceability (0.1)	Free support (0.7)	0	1	1	–
		Type of support (0.3)	24/7, Diagnostic tools, Phone, Urgent response	24/7, Diagnostic tools, Phone, Urgent response	24/7, Phone, Urgent response	24/7, Phone
Cost (0.3)	On-going cost (1)	VM cost (0.6)	$0.68	$0.96	$0.96	<$1/h
		Data (0.2)	10	10	8	10 GB/month
		Storage (0.2)	12	15	15	15 GB
Performance (0.3)	Service response time (1)	Range (0.5)	80–120	520–780	20–200	60–120 s
		Average value (0.5)	100	600	30	–
Security (0.05)	Level: 0–10 (1)	–	4	8	4	4

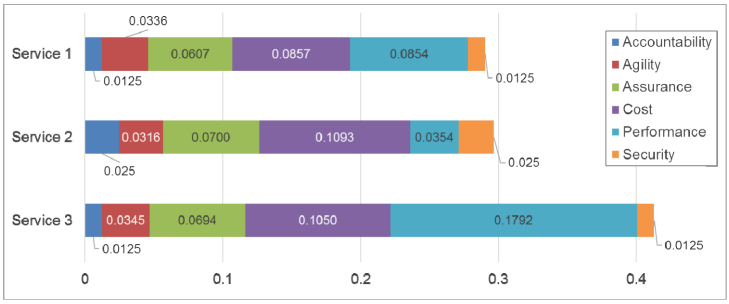


Fig. 2. Ranking in case study example

5 Conclusion

Cloud computing has become important and its paradigm is towards to mobile cloud with the mobile network. Currently, there are many CSPs who offer different services with different quality attributes with their own policy. With the growing number of cloud offerings, there are some of studies for quality assessment of cloud services. However, CSC should be able to verify the QoS and choose an appropriate service provider themselves. However, it is very difficult to compare various services.

In this paper, we proposed a ranking decision model for assessment of cloud services. To design our proposed model, we chose some that are based on KPIs from SMI model made by CSMIC. And we also applied AHP method to decide ranking of cloud services. After that, we showed 4 phase processes how cloud services are assessed through a case study. We expect that our proposed model will help CSCs who choose the service and CSPs operate a cloud service recommend system when reflecting actual user experience.

Acknowledgement. This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2018-2013-1-00717) and the National Program for Excellence in SW (2017-0-00093), supervised by the IITP (Institute for Information & communications Technology Promotion).

References

1. Kim, D.U., Lee, G.H., Kim, Y.S.: Trend of service level agreement in Korea. Electron. Telecommun. Trends **19**(6), 55–65 (2004)
2. Cloud Service Measurement Initiative Consortium: Service Measurement Index Version 1.0 (2011)
3. Giegel, J., Perdue, J.: Cloud service measures for global use: the service measurement index. In: 2012 Annual SRII Global Conference, pp. 411–415. IEEE Press, New York (2012)
4. Kafetzakis, E., Koumaras, H., Koutis, M.A., Koumaras, V.: QoE4CLOUD: a QoE-driven multidimensional framework for cloud environment. In: 2012 International Conference on Telecommunications and Multimedia, pp. 72–82. IEEE Press, New York (2012)
5. Al-Shammari, S., Al-Yasiri, A.: Defining a metric for measuring QoE of SaaS cloud computing. In: 15th Annual Post Graduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting, pp. 251–256. John Moores University, Liverpool (2014)
6. Saaty, T.L.: Theory and Applications of Analytic Network Process: Decision Making with Benefits, Opportunities, Costs, and Risks. RWS Publication (2005)
7. Li, A., Yang, X., Kandula, S., Zhang, M.: CloudCmp: comparing public cloud providers. In: 10th ACM SIGCOMM Conference on Internet measurement, pp. 1–14. ACM, New York (2010)
8. Schad, J., Dittrich, J., Quiané-Ruiz, J.A.: Runtime measurements in the cloud: observing, analyzing, and reducing variance. Proc. VLDB Endowment. **3**(1-2), 460–471 (2010)
9. Iosup, A., Yigitbasi, N., Epema, D.: On the performance variability of production cloud services. In: 11th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, pp. 104–113. IEEE Press, New York (2011)

Design of Science App Management Tool for Computational Science Engineering Platform

Inho Jeon, Yejin Kwon, Jin Ma, Jongsuk Ruth Lee, and Jerry H. Seo(✉)

National Institute of Supercomputing and Networking, Korea Institute of Science and Technology Information (KISTI), Daejeon, Republic of Korea
{inojeon, yejinkwon, majin, jsruthlee, jerry}@kisit.re.kr

Abstract. The development of open science has resulted in the sharing of a greater amount of information including research processes as well as research results. We are developing a platform, EDISON (EDucation-research Integration through Simulation on the Net) to further promote the open science practice in the fields of computational science and engineering. The EDISON platform not only provides an online service of simulations developed by computational scientists and engineers, but also an environment for sharing source codes, data, and related publications. An efficient management system for simulation software programs is required to ensure the service quality of the EDISON platform. This study proposes a registration and management system that introduces simulation software programs developed by computational researchers to EDISON. The developer can register the simulation software programs on EDISON using our system, and establish an effective web-based simulation environment without administrator intervention.

Keywords: Computational science engineering · Web-based simulation · Simulation software · Open science

1 Introduction

In computational science and engineering, researchers should conduct various numerical simulations and visualization experiments to predict and analyze various natural science and engineering problems. This field currently shares a variety of results produced through computational science research including research publications, data, simulation software, and source code in accordance with the current open-science policy. Researchers build independent websites to share simulation software programs, or use open-source repositories such as Github to share source codes. Open Source platforms such as OpenFOAM in computational fluid dynamics [1], SIESTA in nanophysics [2], and nwChem in computational chemistry [3] are the representative online platforms being shared through the Internet.

However, the simulation software are difficult for researchers and students to use, as many operate on the Linux or Unix environments. Further, it is not a simple task to

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 574–579, 2020.
https://doi.org/10.1007/978-981-13-9341-9_99

operate them on PC because they require high computing capacity for parallel processing and an advanced environment for big data processing. Owing to these characteristics, the range of users who are able to benefit from open software is limited even though the resources are shared.

KISTI has been operating EDISON (EDucation-research Integration through Simulation on the Net) since 2011 to increase the accessibility to web-based software in computational science and engineering to vitalize the research community of the fields [4]. The EDISON project provides an online environment for sharing web-based simulation services, source codes, data, and related publications to further promote the practice of open science in computational science and engineering [5, 6].

In this study, we implemented a service for registering and managing simulation software programs on the EDISON platform. Using our system, computational scientists and engineers can register their simulation programs on EDISON. Simulation programs registered on EDISON can utilize computing resources provided by the KISTI, and large-scale research operations that are difficult to conduct on personal computers can be performed. Users can perform simulations and verify the results in any location that has Internet access.

2 Design of the Registration and Management System

The purpose of the science application registration and management system on EDISON is to provide a system that allows computational science and engineering researchers to easily perform registration and management. Table 1 shows the requirements of the registration and management system.

Table 1. Design requirements of science app management tool for computational science engineering platform.

Requirement	Description
Versatility	Simulation codes written in various programming languages such as C++, FORTRAN, and Python should be registered and managed
Expandability	Registered simulation software should be compatible with other programs
Convenience	Technical details that require learning for registration and management should be minimized. The platform should be designed to minimize the source code correction required for registration
Openness	Simulation software should be shared in various forms required by researchers, and should be compatible with open-source repositories
Automation	Simulation software should minimize administrator intervention in the registration process

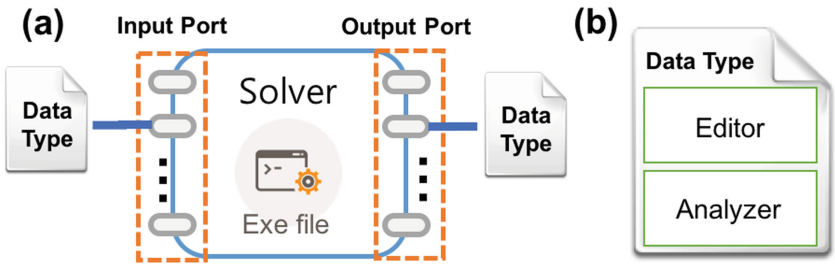


Fig. 1. Simulation components of the computational science engineering platform. (a) Solver (b) Data type

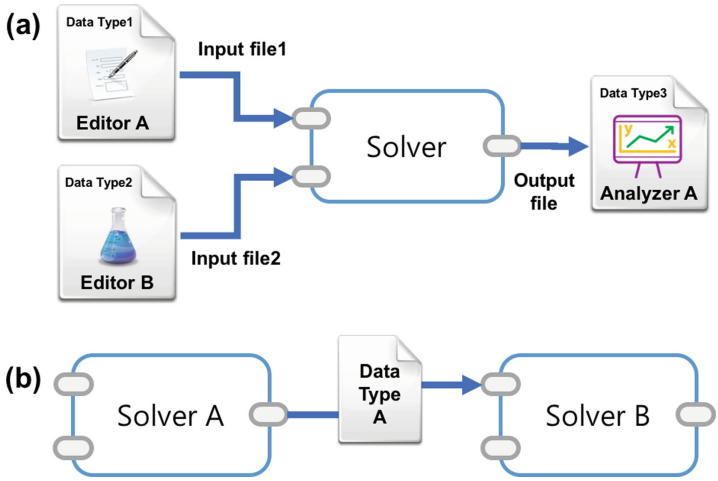


Fig. 2. An execution scenario of (a) Single simulation (b) Workflow simulation

2.1 Designing Simulation Components

A web-based simulation software used to solve the specific problems on EDISON is defined as a science application. In the cases of operating a single simulation and a workflow that executes science applications consecutively, the simulation components are largely categorized into five types that allow for the operation of registered applications. A science application consists of a web editor that processes the input data, a solver that solves problems with the input data, and an analyzer that performs web analyses of the solved result data. The solver has input and output ports to define the input and output forms. It can configure the data type that determines the data form for the input and output ports.

Figure 1 is a diagram showing the structure of the solver and data type. A solver has multiple input and output ports. Each input and output port is designed to have a single data type. The data type has one or more editors and analyzers, and can have connections to multiple input and output ports. A science application developer can use

the existing data type when registering the input and output ports. Figure 2(a) shows a science application with a single solver with two input ports and a single output port. The respective data type was configured at each port of two inputs, and the input files are generated through an editor of the data type. The output port has a data type configured, and the result files can be verified through the analyzer of each data type. Figure 2(b) shows the operation scheme with two interoperable solvers using the workflow [5]. When the data type of the solver A output port and that of the analyzer B input port are the same, the two solvers can be interconnected for a single operation (Table 2).

Table 2. Components for simulation in the computational science engineering platform

Requirement	Description
Solver	A Linux execution file or a script file with input and output ports developed to solve the problems in science and technology
Input/output port	A component that defines the form of input and output files for the solver. It has a single-data type
Data type	A component that classifies the forms of data generated on science applications. It has one or more editors and analyzers
Editor	A component that receives the user's input through the web GUI and generates the input data for the solver
Analyzer	A component that analyzes the result data generated by the solver on the web

2.2 Designing the Science Application Registration and Management System

Figure 3 shows the structure of the science application registration and management system proposed in this paper. Our system offers management functions for the application, data type, editor, and analyzer.

The user can register a new application or manage the existing applications on the application management interface. The system consists of a basic information input stage, a solver information input stage, an input and output port information input stage, and an application testing stage for a new application registration. At the application information input stage, the user can input introductory information including the application title and description. At the stage of entering solver information, the operation method for the execution file is configured, and the execution file is uploaded. The execution files compiled on the developer server provided by EDISON or source code can be uploaded. When uploading the source codes, they are automatically compiled. The program developer can verify errors and results produced from the compiling process.

In the stage of entering the input and output port information, the input and output ports should be defined, and the data type is determined for each port. The data type can be selected from the existing or newly created ones.

Upon data creation, an editor and analyzer can be designated to the new data type. The editor and analyzer also offer registration and management functions, which allow

the registration of science applications of various fields. At the stage of application testing, the application is operated based on the previously entered information. Developers can request for the registration of applications whose testing procedure is completed. The requested applications are made public after being approved by the website administrator, and can then be used from the Science Appstore.

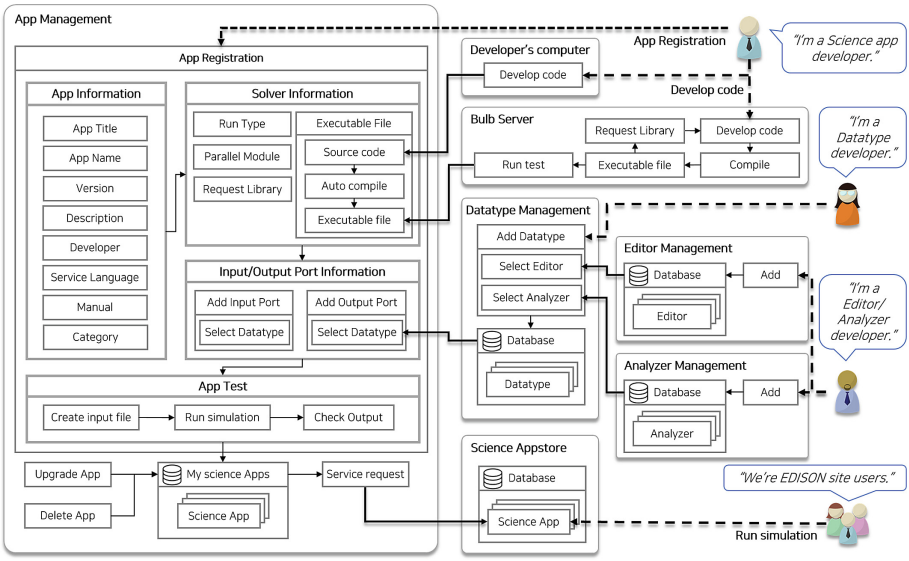


Fig. 3. The architecture of simulation software management system

3 Conclusions

We proposed an online registration and management system for the simulation software in computational science and engineering fields. Our system allows the registering of simulation components including solver, input and output port, data type, editor, and analyzer. It allows computational scientists and engineers to register their simulation software online using our editor and analyzer without expert knowledge in web programming. We established a management system for each simulation component of our registration and management system, and defined the correlation between the components. The solver offers services in various languages including C, Fortran, Python, and R, which guarantees the system versatility. The system ultimately allows users to manage the entire registration process without requiring the intervention of an administrator.

Acknowledgments. 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-18-L12-C06-S01).

References

1. OpenFOAM. The open source CFD toolbox. http://www.openfoam.com. Accessed Aug 2018
2. SIESTA. https://departments.icmab.es/leem/siesta. Accessed Aug 2018
3. nwChem. http://www.nwchem-sw.org. Accessed Aug 2018
4. EDISON. http://www.edison.re.kr. Accessed Aug 2018
5. Suh, Y.-K., et. al.: EDISON: a web-based HPC simulation execution framework for large-scale scientific computing software. In: Proceedings of the ACM/IEEE CCGrid 2016, May 2016, pp. 608–612. https://doi.org/10.1109/CCGrid.2016.31
6. Jeon, I., et al.: Development and application of smart-learning technology in science and engineering fields. KIPS Rev. **23**(1), 16–22 (2016)

Movie Recommendation System Using *k*-clique and Association Rule Mining

Phonexay Vilakone, Doo-Soon Park(✉), and Khamphaphone Xinchang

Department of Computer Sciences and Engineering, Soonchunhyang University, Soonchunhyang-ro 22, Sinchang-myeon, Asan-si, Chungcheongnam-do, South Korea
xayus@yahoo.com, parkds@sch.ac.kr, khamphaphone12@gmail.com

Abstract. Various recommender systems have been presented in an effort to get better preciseness. In order to further improve more accuracy, the *k*-clique methodology, which is used to analyze social networks was introduced in the recommendation system and the result shows the *k*-clique method is effective in improving accuracy. In this paper, we propose a recommendation system using *k*-cliques and association rule mining with the best accuracy. To estimate the performance, the maximal clique method, collaborative filtering methods are monitored using the *k* nearest neighbors, the *k*-clique method, and the *k*-clique and association rule mining are used to evaluate the MovieLens data. The performance outputs show that the *k*-cliques and association rule mining get better the preciseness of the movie recommendation system than any other methods used in this experiment.

Keywords: Association rule mining · Collaborative filtering using a *k* nearest neighbor · *k*-cliques method · Maximal clique method · Movie recommendation system

1 Introduction

Various researchers pay attention to the way communities in complicated networks, for example, the *k*-cliques method introduced in the social network analysis [1–3] is presented in the movie recommendation system as to get better preciseness of the movie recommendation system [4, 5].

The purpose of this article is to get an upwards proficient method than the *k*-clique method. The proposed movie recommendation system is based on the *k*-clique and association rule mining. In the proposed method, the similarity among users is calculated with the help of the cosine similarity measure method. Then, several of the group is created by using a *k*-clique method. Finally, collaborative filtering and association rule mining methods are used to recommend the movies in a group with similarity through the user. The proposed solution gain more effective performance. For performance evaluation, the MovieLens dataset was divided into experimental data and test data. In order to evaluate the performance, we did a comparison of the collaborative

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 580–585, 2020.
https://doi.org/10.1007/978-981-13-9341-9_100

filtering monitors a *k* nearest neighbor, maximal clique, *k*-clique, and *k*-clique and association rule mining methods.

The remaining of the paper is presented as follows: In Sect. 2, related works in the area are presented. In Sect. 3, the proposed method is to explain further. In Sect. 4, the detail of experimental analysis is presented, the results of the experiments are presented and the comparison to evaluate of the performance is presented by the proposed method with the collaborative filtering using a *k*-nearest neighbor and maximal clique and *k*-clique method. Finally, in Sect. 5, the conclusions are conducted and future works are presented.

2 Related Work

Before introducing the proposed method in the next segment. The meaning of this method, along with some theoretical which the method used in this paper are presented below.

2.1 *K*-cliques Community Detection

K-clique is a community detection method and is a complete communication graph with *k* nodes in a complex network, usually with many graphs of different sizes. In general, the value of *k* is bigger than or equivalent to 3. If *k* equals 2, there is the only margin which has less practical meaning [6–8] (Fig. 1).

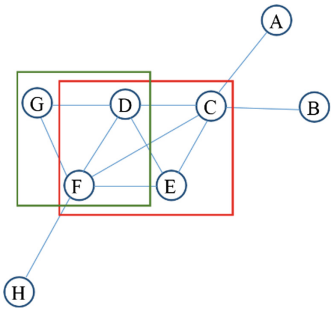


Fig. 1. An example of *k*-clique

Obviously, nodes C, D, E, and F represent groups of 4-clique, because they are connected. Similarly, sets of nodes D, F, and G represent groups of 3-clique.

2.2 Association Rule Mining Method

Association rule mining is the data mining process of finding the rules that may govern associations and causal objects between sets of items. So in a given transaction with multiple items, it tries to find the rules that govern how or why such items are often bought together. For example, peanut butter and jelly are often bought together because

a lot of people like to make PB&J sandwiches. Also surprisingly, diapers and beer are bought together because, as it turns out, that dads are often tasked to do the shopping while the moms are left with the baby [9].

3 Movie Recommendation System Using *k*-Cliques and Association Rule Mining

The process of gathering data and work processes for the guidance system is presented in Fig. 2. Figure 2 presents the movie guide system using *k*-clique and association rule mining.

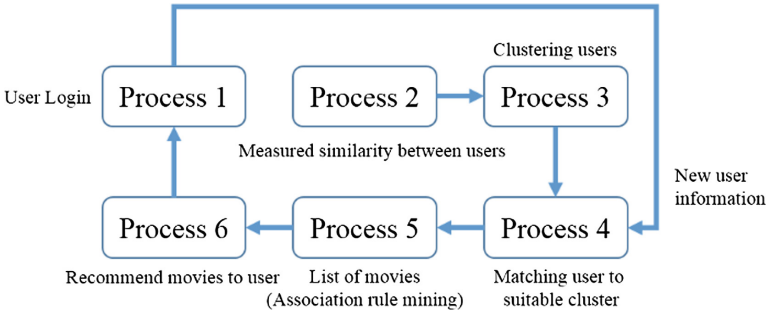


Fig. 2. Flow chart of proposed method.

Process 1, a new user must sign in and provide the necessary personal information, such as gender, age, and occupation.

Process 2, used the experimental data to calculate similarities between users. Similarities in user measurement at the end of this process, an adjacency matrix of user similarities are created.

Process 3, the adjacency matrix is used to divide the user group into several groups with the help of the *k*-clique method. After the completion of the process. Users of multiple groups may present.

Process 4, the user's personal information will be compared to the personal configuration data of the users in each group. The cosine similarity method for calculating the similarity between them. Then the largest group with have number ones will be chosen as the most similar group for the relevant user.

Process 5, the movie arranged by the group members will be sorted by popularity and association rule mining is used to search the movie which all way together viewed.

Process 6, the top 5 movies are often viewed together are introduced to new users.

3.1 Datasets

The proposed method was implemented using the MovieLens dataset [10], this dataset is separated into two datasets. First is experimental data and second is test data. In the

experimental data, there are 800 users, and there are 143 users in the test data. The total rate is 100,000 ratings from 1,684 movies on 943 users. Simple demographics for users, including age, gender, and occupation.

4 Experimental Analysis

After developing the proposed movie recommendation system using *k*-cliques and association rule mining, the number of movies rated by new users in movies recommended by the system was predicted. This paper uses a widely used evaluation metric for benchmarking the output of the proposed method. The mean absolute percentage error (MAPE) is the method of predicting the accuracy of the predictive method in the statistics given by the formula [11–13]:

$$\text{MAPE} = \frac{100\%}{n} \sum_{t}^{n} \left| \frac{A_t - F_t}{A_t} \right| \tag{1}$$

Where A_t is the actual value and F_t is the forecast value.

We compared the MAPE with the maximal clique method, *k*-cliques method, and the collaborative filtering using the k neighboring neighbor method to estimate the effectiveness of the proposed approach. If the number of results is low, it means that our methods are salutary. For the MAPE calculation, the Eq. (1) is used.

First, we measure the MAPE for the proposed method. The MAPE value is 13.98% (see below) when the value $k = 11$.

$$\text{MAPE} = \frac{100\%}{n} \sum_{t}^{n} \left| \frac{A_t - F_t}{A_t} \right| = 13.98\% \tag{2}$$

The MAPE values using the *k*-clique and the association rules are shown in Fig. 3 below. The results of the experiments depend on the values of *k*. Figure 3 shows the minimum values MAPE is 13.98%, with $k = 11$. Figure 3 also shows the results of *k*-clique without association rules. Second, the maximal clique method was calculated the MAPE. The MAPE value was 27.55%. Third, we calculated MAPE for collaborative filtering using k neighboring neighbors. The value of MAPE was 18.88%.

After comparing these methods, the MAPE results calculated by the movie recommendation system using the *k*-clique and association rule mining, a movie recommendation system based on *k*-clique, and that of the movie recommendation system based on maximal clique algorithm, and a movie recommendation system using the collaborative filtering monitored using the k nearest neighbor are more accurate and effective. Figure 4 below, confirm that the performance of four methods at our best.

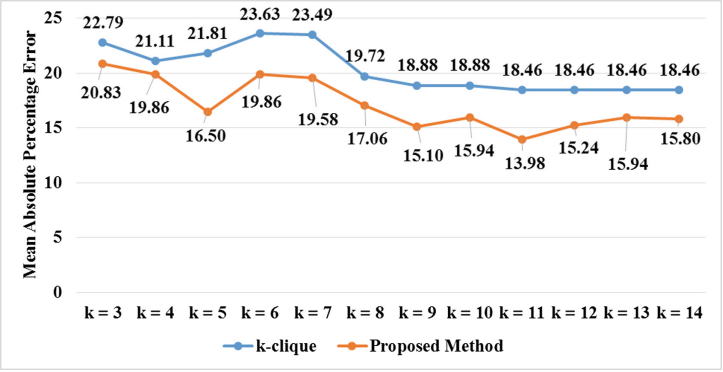


Fig. 3. Result of *k*-clique and proposed method

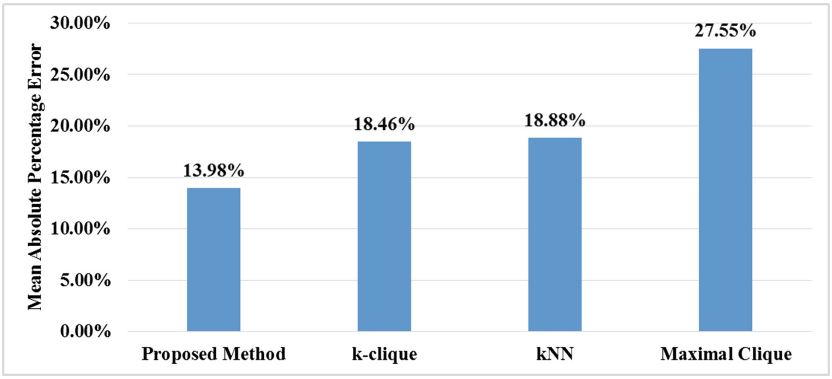


Fig. 4. Comparison of existing method

5 Conclusions

To get more accurate recommendations, the *k*-clique algorithm is very effective in social networks and the association rule mining method, which is used for finding the often together of items in data mining is introduced in this proposed. The result of the experiment, as shown in Fig. 3, the best method was found when $k = 11$.

For output estimation, the *k*-clique method, maximal clique method, collaborative filtering method using a *k* nearest neighbor, and *k*-clique and association rule mining are evaluated. The results showed that the *k*-clique and association rule mining is get more better the preciseness of the movie recommendation system than the other methods used in this.

Future studies can increase the accuracy and effectiveness of the movie recommendation system using *k*-clique. In addition, a Normalized Discounted Cumulative Gain (NDCG) algorism will be used in the future to increase more accuracy with the *k*-clique method.

Acknowledgments. This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2018-2014-1-00720) supervised by the IITP (Institute for Information & communications Technology Promotion) and the National Research Foundation of Korea (No. NRF-2017R1A2B1008421).

References

1. Hao, F., Park, D.S., Pei, Z.: When social computing meets soft computing: opportunities and insights. Hum. Cen. Comput. Inf. Sci. **8**(8), 1–21 (2018)
2. Hao, F., Sim, D.S., Park, D.S., Seo, H.S.: Similarity evaluation between graphs: a formal concept analysis approach. J. Inf. Process. Syst. **13**(5), 1158–1167 (2017)
3. Hao, F., Park, D.S., Min, G., Jeong, Y.S., Park, J.H.: *K*-cliques mining in dynamic social networks based on triadic formal concept analysis. Neurocomputing **209**, 57–66 (2016)
4. Jeong, W.H., Kim, S.J., Park, D.S., Kwak, J.: Performance improvement of a movie recommendation system based on personal propensity and secure collaborative filtering. J. Inf. Process. Syst. **9**(1), 157–172 (2013)
5. Viana, P., Pinto, J.P.: A collaborative approach for semantic time-based video annotation using gamification. Hum. Cen. Comput. Inf. Sci. **7**(13), 1–21 (2017)
6. Hao, F., Park, D.S., Pei, Z.: Detecting bases of maximal cliques in social networks. In: MUE2017, Seoul, Korea, p. 1 (2017)
7. Hao, F., Min, G., Pei, Z., Park, D.S., Yang, L.T.: *K*-clique communities detection in social networks based on formal concept analysis. IEEE Syst. J. **11**, 250–259 (2015). https://doi.org/10.1109/jsyst.2433294
8. Gregori, E., Lenzini, L., Mainardi, S.: Parallel (*k*)-clique community detection on large-scale networks. IEEE Tran. Parallel Distrib. Syst. **24**(8), 1651–1660 (2013)
9. Surbhi, K.S., Jalpa, T.P.: A survey on association rule mining. In: International Conference on Advanced Computing and Communication Technology, IEEE Conference, pp. 212–216 (2015)
10. Harper, F.M., Joseph, A.K.: The MovieLens datasets: history and context. ACM Trans. Interact. Intell. Syst. (TiiS) **5**(4), 19 (2015). Article 19
11. Tofallis: A better measure of relative prediction accuracy for model selection and model estimation. J. Oper. Res. Soc. **66**(8), 1352–1362 (2015)
12. Hyndman, R.J., Anne, B.K.: Another look at measures of forecast accuracy. Int. J. Forecast. **22**(4), 679–688 (2006)
13. Kim, S., Kim, H.Y.: A new metric of absolute percentage error for intermittent demand forecasts. Int. J. Forecast. **32**(3), 669–679 (2016)

QoE Unfairness in Dynamic Adaptive Streaming over HTTP

Geun-Hyung Kim(✉)

Game Engineering Major, Dong-eui University,
176 Eomgwang-no BusanJin-Gu, Busan 47340, Korea
geunkim@deu.ac.kr

Abstract. Today, video streaming over the Internet is predominantly using Hypertext Transfer Protocol (HTTP) as its primary protocol over the Transmission Control Protocol (TCP). To allow heterogeneous devices on the heterogeneous networks to be able to play out the video stream, several HTTP adaptive streaming (HAS) technologies have been employed in numerous video services by adapting the video streaming to the network conditions. Dynamic Adaptive Streaming over HTTP (DASH) has been introduced as an interoperable solution. Since DASH clients may share one or more bottleneck links in a network, such as a home networks and other network links and request the video segments according to the buffer state, the bandwidth unfairness among clients and the underutilization of network bandwidth are inevitable. In this paper, we investigated some reasons why the Quality of Experience (QoE) unfairness occurs.

Keywords: HTTP-based adaptive streaming · MPEG-DASH · TCP flow fairness · QoE unfairness

1 Introduction

A video streaming service should accommodate heterogeneous requirements due to the variety of contents, user preferences and context, devices, and access network constraints etc. To meet these requirements, several video delivery schemes have been developed and can be classified by the type of network management used to establish required bandwidth and to delivery videos. The video streaming services have been implemented over either managed or unmanaged networks [1].

In video streaming over the managed networks, the video service providers should cooperate with Internet service providers (ISPs) that have authority for controlling network resources. In video streaming over the unmanaged networks, video service providers deliver their contents without any technical support from network operators. Today's popular over-to-top (OTT) service is a typical example of unmanaged video

This work was supported by National Research Foundation of Korea (NRF) grant funded by the Korea government (MOE: Ministry of Education): (NRF-2017R1D1A1B03035074).

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 586–591, 2020.
https://doi.org/10.1007/978-981-13-9341-9_101

streaming service. The primary protocol for OTT services is the Hypertext Transport Protocol (HTTP) on top of Transmission Control Protocol (TCP).

Since HTTP can traverse firewalls and NAT and support a progressive download using conventional Web servers and HTML5 enabled browsers, most video streaming solutions are based on HTTP. However, since the network conditions, such as available bandwidth and end-to-end delay, change during a streaming session in realistic networks, the static quality selection may be unsuitable. So, dynamic adaptive streaming over HTTP (DASH) was introduced to adapt on the fly to the network conditions by switching the video quality during the same streaming session.

In DASH streaming, DASH rate adaptation mechanism and TCP congestion control mechanism are conducted at the application layer and the transport layer respectively. These two control mechanisms are unaware of each other and have different goals. The standard TCP congestion control mechanisms aim at equalizing the rate of competing TCP flows in the bottleneck link, regardless required rate of these flows, and the DASH rate adaptation mechanisms aim at preventing video stalls caused by client playback buffer drainage. Due to the limitation of playback buffer size in DASH client (hereinafter referred to as "client"), request patterns from the client may be scattered and have the impact on the performance of TCP variations. Given that DASH clients are in the same access network, several clients may share the bottleneck link and compete each other over the link. This phenomenon makes TCP throughput variation, packet losses, and quality of experience (QoE) unfairness among clients. In this paper, we investigate the impact of end-to-end delay for QoE unfairness among the clients and TCP throughput.

This paper is organized as follows. Section 2 provides related work discussed in this paper. In Sect. 3, we show simulation results to investigate the TCP performance according to the end-to-end delay. Finally, we conclude this paper in Sect. 4.

2 Related Work and Problem Statement

Over the past few years, DASH has widely deployed to support not only VoD services but also live media services through heterogeneous clients and heterogeneous networks thanks to its adaptability to network bandwidth changes. DASH standard only defines two formats; the Media Presentation Description (MPD) format and the segment format (e.g., isobmff, m2ts, etc.). The MPD describes the content characteristics such as content availability, timing, resolutions, media types, and required bandwidths, their URL address, required digital right management (DRM). The segments are the media stream chunks in temporal sequence, consisting of the actual multimedia bit streams and storing in single or multiple files [2].

DASH leverages the existing HTTP based media content delivery infrastructure without any specialized servers. In order to play the media content, the client obtains the MPD file at first. And then it constructs appropriate segment URLs to access segments and selects the appropriate segment typically based on the estimated end-to-end bandwidth and its rendering capabilities. Once playing the content has started, the client continues consuming the media segments and requesting next segments in the chronological order to the Web server.

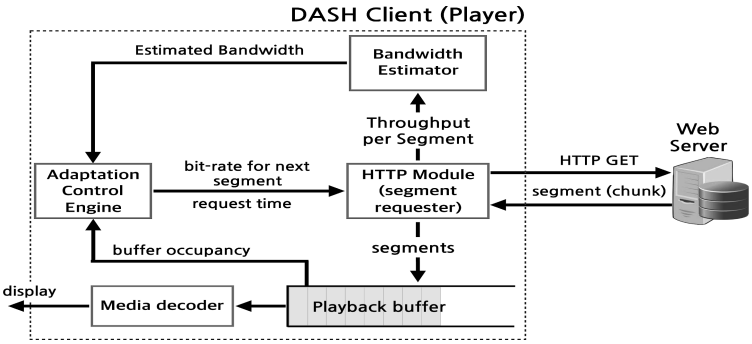


Fig. 1. DASH client architecture for adaptive streaming.

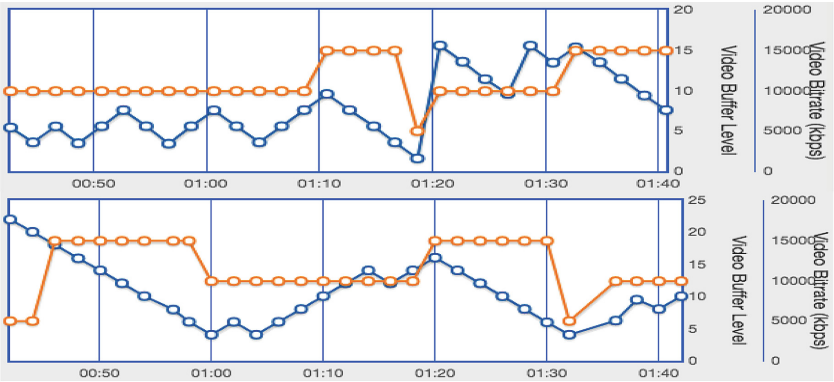
Figure 1 shows the common DASH client architecture, consisting of HTTP module, bandwidth estimator, adaptation control engine, playback buffer and media decoder. Client has two phases alternately: buffering phase and steady phase [3]. The buffering phase is classified into the *initial loading* and the *stalling* depending on when buffering happens without playing. In the buffering phase, client requests a set of segments in the lowest quality level to fill the playback buffer as quickly as possible, and consequently to reduce initial playback delay. Once playback buffer is filled, client enters into the steady phase, during which it requests on and off the segments keeping the fullness of playback buffer.

When family members in the house watch different video contents at the same time, two or more clients might share the residential access link and compete each other for the available bandwidth. As another example of such competition is when many users, in the same edge network, watch the same live events (e.g., Olympic or Super Bowl etc.) over Internet [3]. It has been observed that such competition can lead to performance issues. They also discussed the underlying reason of the performance issues aroused when multiple clients compete.

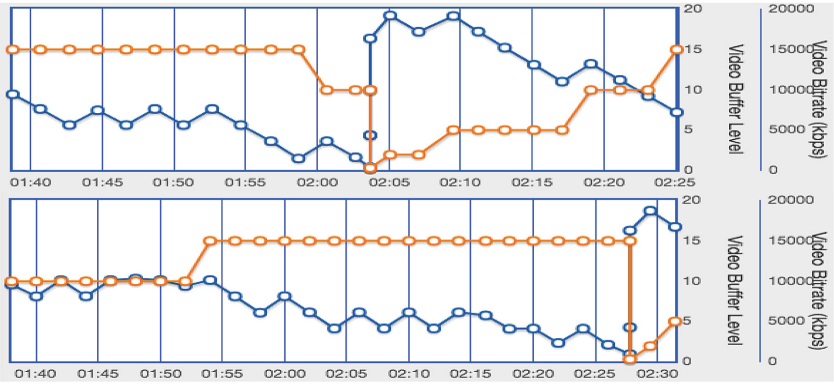
Huang et al. found that the accurate client-side bandwidth estimation above HTTP layer is hard by measurement of three popular HTTP-based video streaming services [4]. They argued that adaptive rate selection algorithms underestimate available bandwidth, because of interaction among the playback buffer, HTTP, and TCP's congestion control algorithm. When the playback buffer fills, on-off scheduling for segment request may change TCP status into slow-start that can lead to incorrect estimation of available bandwidth.

Hu et al. [5] investigated user behaviors in DASH and performance of TCP for DASH by conducting measurement studies on traces of segment requesting and TCP connection. Their study reveals the mutual effect between TCP strategies and segment request patterns in DASH. Request patterns in DASH have a great impact on the performance of TCP and conventional TCPs' strategies may cause user's QoE degradation in DASH. However, they did not consider the impact of end-to-end delay of TCP flows on QoE unfairness.

We investigated the bandwidth unfairness of two clients competing in the shared network by using DASH JavaScript Reference Player [6] released in DASH Industry Forum [7]. In our investigation, we used DASH Reference Player 2.6.4, written using JavaScript, which works in any HTML5 browser supporting the Media Source Extensions and Encrypted Media Extensions. For simplicity, we used two clients (with chrome browsers and dash.js) residing in the same network and one MPD file [8]. We measured the buffer level and video bit rate of two clients 20 times and showed one of those data in Fig. 2 to represent the competing effect of two clients in terms of unfairness. The video segment bit rate requested by each client is different from each other in 2 min playback. From this measurement, we found that the video bit rate and corresponding buffer level changes dynamically and the available bandwidth in the shared network is not fairly allocated for both competing clients.



(a) The buffer level and video bitrate of client 1(up) / client 2 (down)



(a) The buffer level and video bit rate of client 1(up) / client 2(down)

Fig. 2. The measurement in buffer level and video bit rate of competing clients

3 Performance Evaluation

In this paper, we investigate the impact of end-to-end delay difference on TCP performance when the shared link (5 Mbps) is congested and it is not congested. In this simulation, eight TCP flows are classified into two groups that have different end-to-end delay. Node 1–4 are in the same link whose delay varies from 2 ms to 20 ms and node 5–8 are in the same link whose delay is 2 ms. The delay differences are 0 ms, 2 ms, 8 ms, 13 ms, 18 ms. When there is no congestion in the shared link, the difference in end-to-end delay has no any effect on the throughput (left in Fig. 3). On contrary, the end-to-end delay has effect of the throughput when there is congestion in the shared link (right in Fig. 3). As shown in right figure in Fig. 3, the throughput of nodes in the slow link is close to zero when the end-to-end delay is greater than 13 ms.

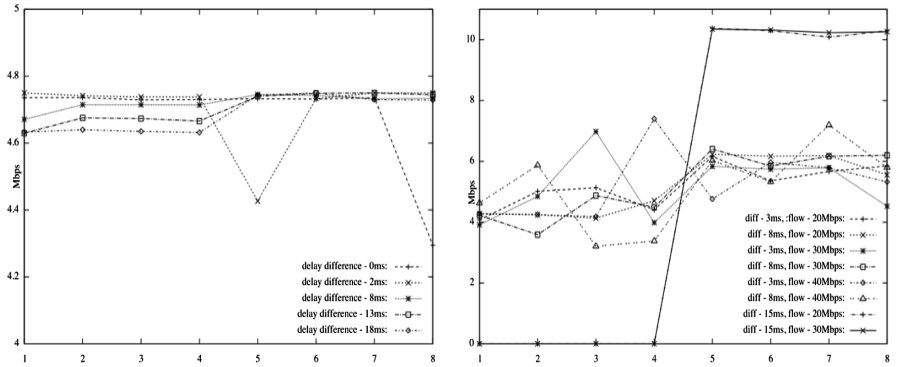


Fig. 3. TCP throughputs w/o or w/congestion (left: without congestion, right: with congestion)

4 Conclusion and Future Work

In this paper, we have investigated the reason why the QoE unfairness occurs, by measuring the buffer level and video rate of DASH reference clients in the real network and simulating TCP performance and DASH clients to find the effect of end-to-end delay difference in connections.

As the future work, we consider two activities. At first we will implement the proposed mechanism in the DASH reference client (dash.js) and verify it in the real environment. In addition, we will investigate the performance issues of the DASH over QUIC (Quick UDP Internet Connection) instead of the DASH over TCP.

References

1. Begen, A., Akgul, T., Baugher, M.: Watching video over the web: part 1: streaming protocols. Internet Comput. IEEE **15**(2), 5463 (2011)

2. Sodagar, I.: White paper on MPEG-DASH Standard, ISO/IEC/JTC1/SC29/WG11W13533, April 2012
3. Akhshabi, S., Anantakrishnan, L., Dovrolis, C., Begen, A.C.: What happen when HTTP adaptation streaming players compete for bandwidth? In: Proceedings of the 22nd International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV), Toronto, pp. 9–14. ACM (2012)
4. Huang, T.-Y., Handigol, N., Heller, B., McKeown, N., Johari, R.: Confused, timid, and unstable: picking a video streaming rate is hard. In: Proceedings of the 2012 Internet Measurement Conference (IMC), Boston, pp. 225–238. ACM (2012)
5. Hu, W., Wang, Z., Sun, L.: A measurement study of TCP performance for chunk delivery in DASH, arXiv preprint arXiv:1607.01172 [cs.MM] (2016)
6. DASH JavaScript Player (2018). https://reference.dashif.org/dash.js/v2.6.4/samples/dash-if-reference-player/index.html
7. DASH Industry Forum (2018). http://www.dashif.org
8. Reference MPD file (2018). https://dash.akamaized.net/akamai/bbb_30fps/bbb_30fps.mpd

Interference-Aware Routing for Multi-hop Energy-Constrained Wireless Network with SWIPT

Shiming He[1(✉)], Kun Xie[2], Jin Wang[1], and Dafang Zhang[2]

[1] School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha 410114, China
{smhe_cs, jinwang}@csust.edu.cn
[2] College of Computer Science and Electronics Engineering, Hunan University, Changsha 410082, China

Abstract. Simultaneous wireless information and power transfer (SWIPT) can prolong the life of wireless nodes by harvesting energy and decoding information from a same RF signal, which changes the design of energy-constrained wireless network. In multi-hop wireless network, the interference generally exists. Route, interference and SWIPT are dependent. However, the exist works consider SWIPT link resource allocation with given route or only select path for only one flow without interference. Therefore, this paper analyze the influence of interference on SWIPT, and formulate the interference-aware SWIPT routing problem. The simulation results show that it can obtain more performance gains from interference and SWIPT with more flows.

Keywords: Simultaneous wireless information and power transfer · Interference · Multi-hop energy-constrained wireless network · Routing algorithm

1 Introduction

In wireless network, with the help of the simultaneous wireless information and power transfer (SWIPT) [1], the receiver can harvest energy and decode information from the same RF signal transmitted by a sender. SWIPT can prolong the life of wireless nodes, which changes the design of energy-constrained wireless network. After applying the SWIPT in one-hop [2] or two-hop [3] wireless networks, how to exploit SWIPT to improve the performance of multi-hop wireless network has attracted interests [4, 5]. [4] points that with SWIPT, the remaining energy of nodes are more and the distribution of remaining energy can be more balanced.

In multi-hop wireless network, the interference generally exists and the interference makes different influence on SWIPT compared to wireless information transmission (WIT). Route, interference and SWIPT interact and are dependent. However, [4] allocates the SWIPT link resource with given route, that is, the path selection and SWIPT link resource allocation are separated. [5] designs SWIPT routing for only single flow without interference.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 592–598, 2020.
https://doi.org/10.1007/978-981-13-9341-9_102

Therefore, this paper joint considers interference, SWIPT and routing to formulate interference-aware SWIPT routing. We firstly analyze the influence of interference on SWIPT and design a interference based information and energy allocation model to maximize the link capacity with SWIPT. Then we design a interference-aware route metric and formulate interference-aware SWIPT routing problem. The simulation results show that as the number of flows increases, there is more likely to obtain performance gains from interference and SWIPT.

2 System Model

A multi-hop energy-constrained wireless network with N nodes is considered, as shown in Fig. 1. All nodes are equipped with one antenna. A node i's residual energy is denoted E_i. If the node i's residual energy E_i is lower than E_{min}, the node i will refuse to forward information for other nodes in order to prolong it's own life time. The nodes are equipped with SWIPT circuit and can communicate with each other by SWIPT. There are multiple concurrent flows, denoted by a set $F = \{F_1, F_2, \ldots, F_M\}$ of M flows. The data for each flow may traverse multiple hops in the network. A flow $F_i(s_i \rightarrow d_i)$ goes through a pair of source node and destination node, denoted as s_i and d_i, respectively. We want to find path for the k-th flow under the $k-1$ flows information.

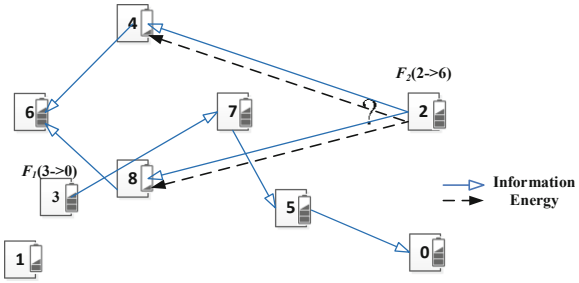


Fig. 1. Network model.

The transmitted signal by the sender node is denoted by $x(t)$ with bandwidth of W Hz. The transmitted signal propagates through a wireless channel with channel gain $|h_{ij}|^2$ which captures the effects of path-loss, shadowing, and fading within the channel. Because of the other flows, there are interference from other nodes. $x_l(t)$ denotes the interference signal from node l with power P_l, and h_{lj} denotes the channel between l and j. The received RF signal at the receiver is $y(t)$

$$y(t) = \sqrt{P_{ij}} h_{ij} x(t) + \sum_{l \in \Phi_j} \sqrt{P_l} h_{lj} x_l(t) + n_{ij}(t) \tag{1}$$

where Φ_j is the node set interfered with node j, n_{ij} is the antenna noise which is circularly symmetric complex Gaussian (CSCG) distribution with mean 0 and variance σ_{ij}^2. If the receiver node is a WIT receiver, the signal-to-noise ratio (SNR) and capacity are given by (2) and (3), as shown in Fig. 2.

$$\gamma_{ij}^{IIT} = |h_{ij}|^2 P_{ij} / (\sigma_{ij}^2 + \eta_{ij}^2 + \sum_{l \in \Phi_j} P_l |h_{lj}|^2) \tag{2}$$

$$C_{ij}^{WIT} = W * \log_2(1 + \gamma_{ij}^{IIT}) \tag{3}$$

where $z_{ij}(t)$ is the additional baseband noise and CSCG distribution with mean 0 and variance η_{ij}^2.

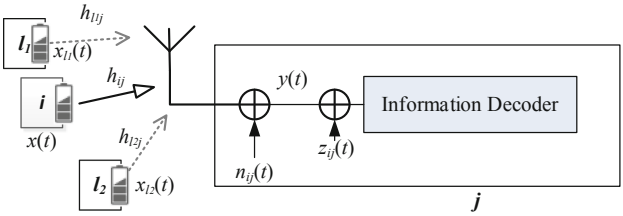


Fig. 2. Wireless information receiver

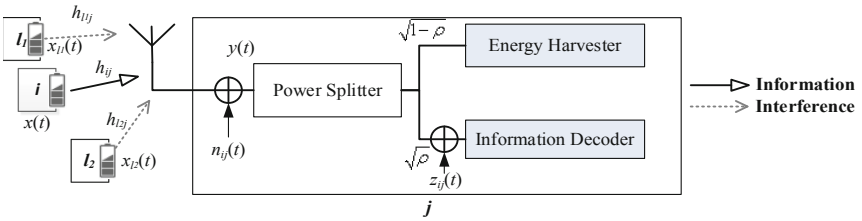


Fig. 3. SWIPT receiver

If the receiver node is a SWIPT receiver, the received RF signal is divided into two parts with different power by the power splitter, as shown in Fig. 3. The a power splitting ratio is denoted by $\rho \in [0, 1]$. Therefore, the RF signal $\sqrt{1-\rho}y(t)$ is fed into the energy harvesting circuit and the other RF signal $\sqrt{\rho}y(t)$ is fed into the signal processing circuit. Therefore, the signal for information decoding (ID) and energy harvesting (EH) are given by (4).

$$\begin{aligned} y^{ID}(t) &= \sqrt{\rho}y(t) + z_{ij}(t) \\ y^{EH}(t) &= \sqrt{1-\rho}y(t) \end{aligned} \tag{4}$$

The SNR and the harvested power are given by (5).

$$\begin{aligned} \gamma_{ij}^{ISWIPT} &= \rho_{ij} |\, h_{ij} \,|^2 P_{ij} / (\rho_{ij}\sigma_{ij}^2 + \eta_{ij}^2 + \rho_{ij} \sum_{l \in \Phi_j} P_l |\, h_{ij} \,|^2) \\ E_{ij}^{leh} &= \varepsilon(1-\rho_{ij})(|\, h_{ij} \,|^2 P_{ij} + \sigma_{ij}^2 + \sum_{l \in \Phi_j} P_l |\, h_{lj} \,|^2) \end{aligned} \tag{5}$$

where ε is the energy converting coefficient of EH circuit. The link capacity with SWIPT is given by (6).

$$C_{ij}^{\text{SWIPT}} = W * \log_2(1 + \gamma_{ij}^{ISWIPT}) \tag{6}$$

If the power splitting ratio is 1, the SWIPT receiver equals to a WIT receiver. Therefore, the SWIPT receiver can also use wireless information transmission.

3 Interference Based Information and Energy Allocation Model

In order to maximize the performance of flow, we first need to maximize the link capacity with interference. Therefore, the information and energy allocation problem to maximize the link capacity with interference can be described as follow:

$$\begin{aligned} &\max_{\rho_{ij}, P_{ij}} C_{ij}^{\text{SWIPT}} \\ &s.t.\ (5)(6),\ E_{ij}^{leh} \geq Pc_j,\ P_{ij} \in [0, P_{\max}],\ \rho_{ij} \in [0,1] \end{aligned} \tag{7}$$

There are three constraints. The harvested power is no lower than the energy harvesting requirement Pc_j, which is dependent on the residual energy and forwarding consume of node j. The transmission power is no larger than the maximum transmission power P_{max}. The power splitting ratio is in the range of 0 to 1. The parameters are power splitting ratio and transmission power. The allocation objective is to maximum the transmission capacity.

According to the model, we can investigate how the interference impacts the link capacity with SWIPT. Given the sender node, receiver node, and interference nodes, we calculate the maximum link capacity with SWIPT by solve problem (7). At the same time, we give the link capacity with WIT by Eqs. (2) and (3) at the same interference. We vary the interference power from 0 mw to 100 mw.

As shown in Fig. 4, when the interference power increases, the link capacities of SWIPT and WIT both become lower. In the first phase, the SWIPT link does not exist, when the interference power is lower than 16 mw. In this case, the received power can't satisfy the energy harvesting requirement. In the second phase, when the interference power is larger than 16 mw, the received power can support the energy harvesting requirement and the SWIPT link exists. At this phase, the link capacity of SWIPT and IT are equal.

Noted that, interference can help build a SWIPT link. Once the SWIPT link is built, the lower interference is better. Therefore, for selecting path with interference, when the SWIPT link doesn't exist, selecting the node with higher interference as the next hop to make sure the connectivity. When the SWIPT link is built, selecting the node with lower interference as the next hop to improve the capacity.

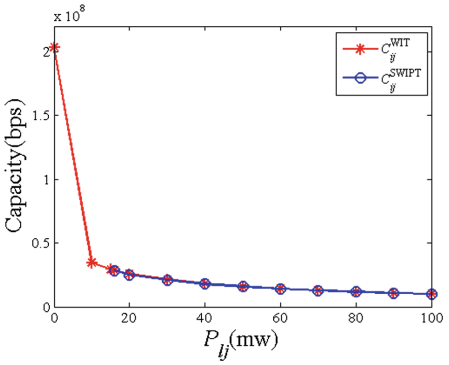


Fig. 4. Capacity of WIT and SWIPT

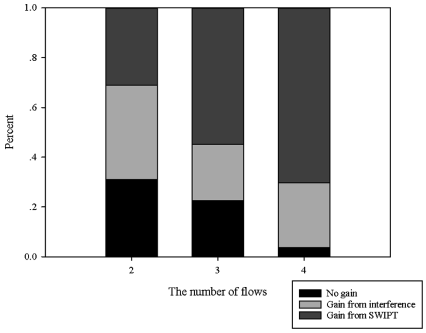


Fig. 5. The gain distribution of flows

4 Interference-Aware SWIPT Routing

Therefore, according to the interference based information and energy allocation model, we design the routing metric and formulate routing problem.

In multi-hop energy-constrained wireless networks, there are two transmission modes (WIT and SWIPT). Based on Eqs. (3) and (6), the routing metric of link (Inference-aware Capacity Available metric, IaCA) is defined as the maximum link capacity among all transmission modes.

$$\mathrm{IaCA}_{ij} = \max\{C_{ij}^{WIT}, C_{ij}^{SWIPT}\} \tag{8}$$

And the routing metric of a path is the minimum of all links' metric in the path.

$$\mathrm{IaCA}_{sd} = \min\{\mathrm{IaCA}_{sc}, \mathrm{IaCA}_{cd}\}, \mathrm{c} \in \mathrm{Path}_{sd} \tag{9}$$

where node c is the node in the path from s to d.

According to the allocation model and flow conservation, the problem of selecting path from s_k to d_k. can be formulated as problem (10).

$$
\begin{aligned}
&\max_{r,P,\rho} \min r_{ij}\mathrm{IaCA}_{ij} \\
&\text{s.t.C1:(8)(2)(3)(5)(6)} \\
&\text{C2:}E_{ij}^{leh} \geq Pc_j, \text{if } E_j < E_{\min}, \forall i,j \\
&\text{C3:}Pc_j = P_{jk}, if\ r_{jk} = 1, \forall j \\
&\text{C4:} \sum_j r_{ij} - \sum_j r_{ji} = \begin{cases} 1, i = s_k. \\ -1, i = d_k. \\ 0, other. \end{cases}, \forall i \\
&\text{C5:} 0 \leq P_{ij} \leq P_{\max}, \forall i,j \\
&\text{C6:}\rho_{ij} \in [0,1], \forall i,j \\
&\text{C7:}r_{ij} \in \{0,1\}, \forall i,j \\
&\text{C8:}i,j \in [1 \ldots N]
\end{aligned}
\tag{10}
$$

where a binary variable r_{ij}, which has value 1 if the link l_{ij} is active in the path of flow, and value 0 otherwise. In C1 constrain, Eq. (8) is the routing metric, Eqs. (2) and (3) are the capacity of link with WIT, and Eqs. (5) and (6) are the capacity of link with SWIPT. C2–C3 constrains are the value of the energy harvesting requirement, which is the power cost for forwarding by node j and set to the transmission power from node j to its next-hop node k. C4 constrain is the flow conservation to find a path from s to d. C5–C8 constrains are the range of parameters. We can solve problem (10) to get the route.

5 Simulation

We take the Fig. 1 as the simulation network topology, and the remaining energy of the 1, 3, 4, and 8 nodes is insufficient to forward data and need to replenish energy. We vary the number of flows from two to four, and randomly select 50 source and destination nodes pairs for each the number of flows. The metric of the evaluation is the path capacity of the last flow, that is, the route metric of the last flow's path.

We evaluate the effectiveness of our scheme by comparing with two different schemes. The first is WIT without interference (WITwoi) which uses the Eq. (3) without the interference part as the route metric. The second is WIT with interference (WITwi) which uses the Eq. (3) as the route metric. The last is our scheme (SWIPTwi).

Previous literature [5] pointed out that according to the distribution of source and destination nodes of the flow, not all flows can take SWIPT to improve performance. Therefore, we analyze the gains distributions of the 50 different flows. The flows can be divided into three categories. In the first category, the path capacities of the all schemes are equal, because the selected paths are same. In the second category, the path capacities of the schemes considering interference are higher than that of the schemes without interference, but the performance of WITwi equals to that of SWIPTwi. In the third category, the performance of SWIPTwi is higher than that of WITwi. It can be seen that the performance is improved by SWIPT. Therefore, these three categories are named as no gain, from interference, and from SWIPT.

Figure 5 shows the proportion of the three categories in all flows. It can be seen that the sum of the third category and the second category is greater than 65% in each the number of flows, and the proportion of the first category decreases significantly as the number of flows increases, and the third category increases from 30% to 70%. This means that as the number of flows increases, there is more likely to obtain performance gains from interference and SWIPT.

6 Conclusion

By joint considering interference, SWIPT and routing, we analyze the influence of interference on SWIPT, design a interference based information and energy allocation model, and formulate interference-aware SWIPT routing problem, which can obtain performance gains from interference and SWIPT. The future work is design algorithm for the routing and taking more experiments.

Acknowledgments. This work was supported by National Natural Science Foundation of China (Nos. 61802030, 61572184, 61502054), the Science and Technology Projects of Hunan Province (No. 2016JC2075), the Research Foundation of Education Bureau of Hunan Province, China (Nos. 16C0047, 16B085).

References

1. Varshney, L.R.: Transporting information and energy simultaneously. In: IEEE International Symposium on Information Theory (ISIT), pp. 1612–1616 (2008)
2. Liu, L., Zhang, R., Chua, K.C.: Wireless information and power transfer: a dynamic power splitting approach. IEEE Trans. Commun. **61**(9), 3990–4001 (2013)
3. Liu, Y., Wang, X.: Information and energy cooperation in OFDM relaying: protocols and optimization. IEEE Trans. Veh. Technol. **65**(7), 5088–5098 (2016)
4. Guo, S., Shi, Y., Yang, Y., Xiao, B.: Energy efficiency maximization in mobile wireless energy harvesting sensor networks. IEEE Trans. Mob. Comput. **17**, 1524–1537 (2017)
5. He, S., Xie, K., Chen, W., Zhang, D., Wen, J.: Energy-aware routing for SWIPT in multi-hop energy-constrained wireless network. IEEE Access **6**, 17996–18008 (2018)

Developing Interactive Planning Methodology for EPC Projects

Jaehyun Choi(✉) and Hankyeom Wang

Korea University of Technology and Education (Koreatech), Architectural Engineering, 1600 Chungjeol-ro, Dongnam-gu, Cheonan-si, Chungcheongnam-do, Republic of Korea
jay.choi@koreatech.ac.kr

Abstract. EPC plant projects are large and complex, requiring systematic working methodologies, accumulated data, and thorough planning through communication between the entities. In this study, the method of extracting the process planning factor information using the asset data of the plant project and using it to present the initial process plan is presented through the concept of IAP (Interactive Planning).

Keywords: Interactive Planning · Work package · Asset data · EPC plant construction

1 Introduction

1.1 Research Background and Purpose

Plant projects are relatively large and difficult, and it is hard to predict and manage due to complex interactions between work species. In order to effectively plan complex and diverse plant projects, it is necessary to define the subjects to carry out the plan, to extract various kinds of data that are the basis of the planning work, and to systematically use them to process accurate and usable data (Hinze 2004). In particular, processes that are closely linked to all the management areas of a project require systematic work through tools to create a process plan, which ensures that accurate and efficient process planning can be used to predict and control the project more accurately (Flood 2007). However, the process plan at the initial stage of the project is at the general level of the milestone level, and this level of process plan does not include important factors in the process planning such as the plant module and major work. In addition, it is not possible to define the detail level of the WBS and organize the activities in the data-limited order stage without specific plans and drawings, and it is difficult to flexibly respond to the changes that occur as the scope of the future project becomes concrete (CII 2013; Pellegrino 2017).

This study suggests a methodology to implement systematic early stage process plan using similar performance data collected from EPC plant project according to this necessity and evaluates and improves process plan through mutual plan, Propose a process planning methodology to ensure that there is no disruption to proceeding. For this purpose, the project participants have extracted the process data as the work

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 599–605, 2020.
https://doi.org/10.1007/978-981-13-9341-9_103

package of the output unit and set the relationship of the extracted work package, and suggests ways to identify predictive and critical management elements.

1.2 Research Scope and Method

In this study, IAP process is defined as core concept and it is structured in accordance with plant project, and IAP's process output is also defined in terms of process management of plant project. IAP is limited to the EPC project in which the subjects are defined to some extent at the early stage of the project because the various fields participating in the project and the participant in the project are carried out at the early stage of the project.

The research method first investigated the management status of the domestic and overseas plant projects, searched the improvement plan based on the project process management, defined the IAP methodology applied to the EPC plant project step by step. In addition, IAP procedures are theoretically defined and applied to sample process scenarios using process data from actual EPC plant projects to verify them.

2 Initial Process Planning Method Using IAP

In order to extract the above process elements using asset process data, we first identify the units to be installed on the site, extract the types to be included in the project, and define the management unit package at the level immediately below the 'discipline' Should be. Each package shall be based on one or more deliverables for each plant module, area, and structure and the essential activities necessary for this shall be included in the package. The package structure of the WBS by work type and the high level formal work according to it can improve the usability of the data by extracting the typical activity data and constituting it as a bundle of units that can be managed in the initial stage; we can continue to improve it. For the systematic implementation of IAP, this study presents three stages and defines each step-level output to finally enable continuous mutual planning during the whole life cycle.

2.1 Process Element Extraction Stage

Process elements are the data necessary for the process planning, including the level-specific contents of WBS and activity information pertaining to them. The main contents included in the WBS are phase, area, discipline, module, unit and can be modified and supplemented to suit individual project. Since discipline is a key element in most plant projects, this study also extracts process elements from the discipline and constructs a work package. In this step, basic information of the project is reviewed and shared, and elements necessary for the execution project process are extracted from the asset data.

2.2 Process Planning Step

Once the necessary work types, packages, and work orders are available for the process plan, you need to define the required packages for the plant module and define the required or subsequent work for each typical activity included in the package. Since all tasks and packages have a line-tracing relationship with other tasks and packages, you can create a unit-specific process table by establishing the relationship between each package and task by linking them. The subject of the process plan can select the work package required by the plant unit and connect them to the process plan. Finally, the final project process plan can be established by establishing the unit relation.

2.3 Process Planning Evaluation Stage

The initial milestones generated during the previous phase must continue to be evaluated and refined until finalized with the project milestones. In particular, the identification of key tasks that can impact the project is the basis for identifying risks and establishing countermeasures.

At this stage, various process plan evaluation criteria are selected according to the characteristics of the project, and work types are evaluated according to the criteria. In order to consider the effect of the work belonging to each type of work on the process plan, critical management factors are extracted through specific criteria. The criteria for extracting tasks are interrelated tasks, long and difficult tasks for procurement, tasks that are continuously required for many structures and units, and long-term tasks for unit work. Standards can be applied.

The method of deriving the key management factor by type of work presented in this study is a method of numerically representing the result of evaluation of several works included in the project process plan according to the number of tasks included in each standard. This evaluation method enables the relative evaluation of the tasks included in the process plan according to the individual characteristics of the project, and it is possible to grasp the elements to be managed with emphasis on each work type (Seo 2017). In addition, because the tasks included in each project are different, the result values can be varied even if the same key management factor derivation criteria are applied, and it is expected that the various characteristics of the work type can be measured by various criteria

3 Applying IAP to the Sample Project Data

In order to apply and verify the IAP methodology proposed in this study, it is ideal to use actual plant project cases (Steiner et al. 2017). However, it is difficult to obtain detailed process data of EPC companies. Therefore, verification using virtual plant project scenarios was performed. The virtual sample project uses the case of the petrochemical plant among various plant types. The work package for each type of product is limited to the most important construction, and the process unit installed in the refinery plant and the water treatment, power supply, storage unit, and control building are selected for the chemical plant (Table 1).

Table 1. Units and disciplines included in the sample project.

Title	Contents
Installed Unit	Process Unit
	Control Building
	Storage Unit
	Power Supplier
	Water Treatment Unit
Discipline	Civil
	Architecture
	Piping
	Electrical
	Instrument
	Steel Structure
	Mechanical
	Painting

3.1 IAP 1st Step: Extracting Project Schedule Element from Asset Data

Application The sample project consists of five major units and several main types of work. Each unit-specific planner extracts the work required for the construction of the unit from the asset data according to the characteristics and constructs the core tasks that follow. In this case, the extracted work is bundled into a product unit package by product type, where the product unit is a sub-discipline unit belonging to the lower stage of the general WBS and the corresponding level unit, and is constructed considering the actual construction.

3.2 IAP 2nd Step: Creating Schedule with Work Packages

The work packages extracted from the asset data and configured are optionally included in the units to be installed in the sample project to constitute a process table for each unit. The milestones that will be created in the sample project are organized into a total of five units, Fig. 1 showing the example of packages required and the process plan for Control Building.

3.3 IAP 3rd Step: Deriving Critical Management Elements Through Schedule Evaluation

In order to evaluate the initial process plan of each case, the evaluation criteria considering the early stage of the project are as shown in Fig. (1) work with a lot of relation with other work, (2) the period of procurement and work with high degree of difficulty, (3) and (4) the period of unit work is relatively long. At that stage, work-specific tasks are evaluated through the frequencies included in each criterion to identify those that may affect future processes.

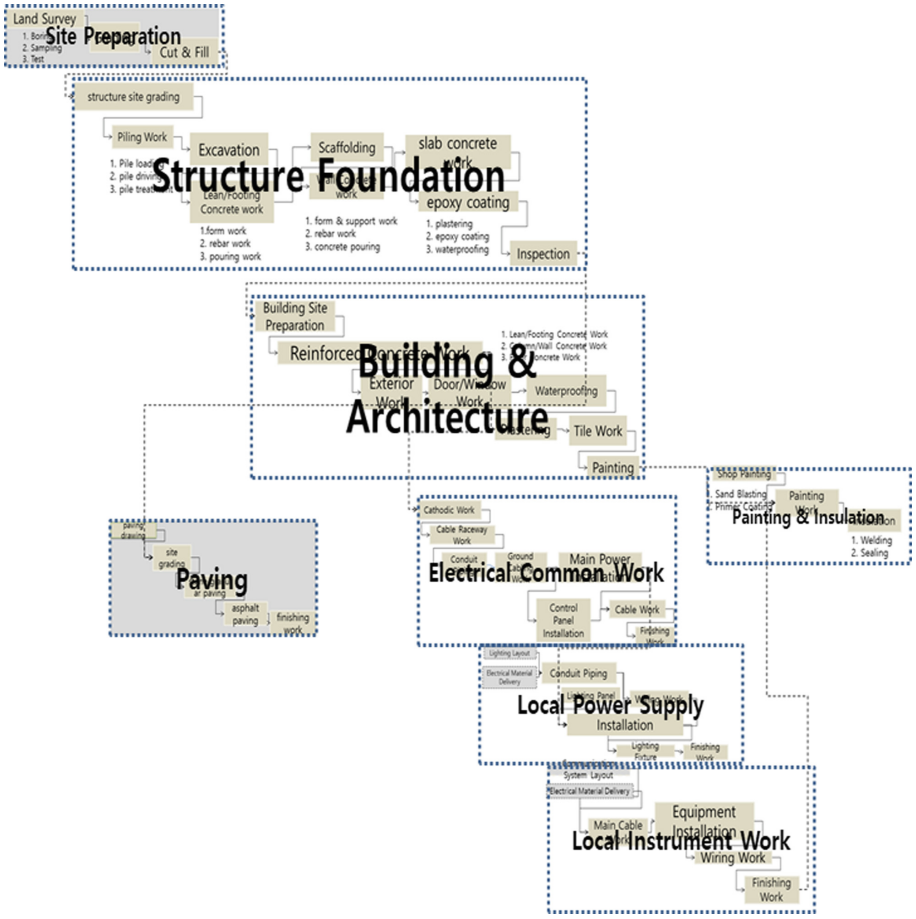


Fig. 1. Extracted work packages and process plan of control building

In order to evaluate the tasks derived by each element by work type, the most frequent value is set to 1 and the remaining values are evaluated as relative values according to the number of elements of the work package and tasks. For example, when the tasks that require Long Lead Items are included in the Mechanical workload, the total number of tasks belonging to Mechanical is set to 1, and the number of tasks belonging to the remaining workloads is set to 1 of the total. The Table 2 shows the number of jobs belonging to the criteria for each type of work and the corresponding conversion values.

Table 2. Number of tasks per work category that fall under the derivation criteria.

Classification	Activity interaction	LLI equipment procurement	Units included frequency	Long duration
Civil	17(1)	0(0)	7(0.44)	9(1)
Architecture	4(0.24)	0(0)	1(0.06)	1(0.11)
Steel Structure	8(0.47)	0(0)	3(0.19)	0(0)
Mechanical	8(0.47)	5(1)	3(0.19)	0(0)
Electrical	11(0.65)	2(0.4)	16(1)	3(0.33)
Piping	15(0.88)	0(0)	7(0.44)	1(0.11)
Instrument	5(0.29)	0(0)	5(0.31)	1(0.11)
Painting	5(0.29)	0(0)	4(0.25)	0(0)

4 Conclusion

Based on the project planning methodology presented in this study, the following conclusions and improvements were derived. First, there is a need for a high-level method to continuously capitalize various information generated during a project in order to carry out IAP more effectively. Secondly, it is not based on the frequency and number belonging to the Critical Management Elements Evaluation Criteria suggested by this study, but by using various objective and scientific evaluation tools, to be evaluated through accurate quantification probability. Third, a more effective IAP can be implemented if it is combined with system tools such as project management information system (PMIS) so that project planning methodology through IAP can be effectively controlled and managed by all project subjects. Fourth, it is necessary to define and document the business process, input and output data, decision makings, roles and responsibilities for each process element of the management unit based on the result of this study.

Acknowledgments. This research was supported by a grant from Construction Technology Innovation Program (18IFIP-B091004-05) funded by Ministry of Land, Transportation and Maritime Affairs (MLTM) of Korean government.

References

Construction Industry Institute: Construction work packages best practice, CWP best practice report, Alberta, pp. 1–8 (2013)

Flood, I.: Project planning using an interactive, structured modeling environment. In: Proceedings of the 2007 Winter Simulation Conference, Florida, pp. 2019–2027 (2007)

Hinze, J.: Construction Planning and Scheduling, pp. 2019–2027. Pearson Education, Florida (2004)

Pellegrino, S.: Introduction to CII's advanced work packaging – an industry best practice. In: Long International, Colorado, pp. 1–22 (2017)

Seo, J.: The development of bidding phase risk response system model for EPC power plant projects, Soongsil University Graduate School, Department of Construction, Seoul, pp. 1–181 (2017)

Steiner, B., Mousavian, E., Saradj, F.M., Wimmer, M., Musialski, P.: Intergrated structural-architectural design for interactive planning. In: Computer Graphics Forum (2017)

Movie Recommendation System Using Social Network Analysis and *k*-Nearest Neighbor

Khamphaphone Xinchang, Doo-Soon Park(✉), and Phonexay Vilakone

Computer Science and Engineering, Soonchunhyang University, Asan, Chungnam, Korea
khamphaphonel2@gmail.com, parkds@sch.ac.kr, xayus@yahoo.com

Abstract. Many types of research have been conducted in recommendation system to develop approaches to solve the challenges for collaborative filtering problem such as cold start problem; in this paper, we proposed the approach to solving the problem of collaborative by using social network analysis and *k*-nearest neighbor (*k*-NN). We used the centrality of social network to detect the community or cluster group to the user, and then apply the *k*-NN method to find a group for new users with similar personal information such as age, gender, and occupation after that recommendation system will recommend items that users in the group were previously interested for the new.

Keywords: Recommendation system · Social network analysis · *k*-Nearest neighbor

1 Introduction

Now a days, the amount of information in the word increases more rapidly than our ability to process. The recommender system is common for e-commerce business to implement. These recommendation systems appeared as a tool to filter information to users based users taste. Recommender systems are used in various areas including book, music, movie, news, CDs, DVDs, TV programs and products in general [1]. There are many ways to build a recommendation system, including providing a list of the top items, building recommendations based on demographics and building suggestion by analyzing various historical preference of users [2]. According to their different ways of producing recommendations, recommender systems are usually divided into three categories: rule-based recommendation system, the content-based recommendation system and collaborative filtering (CF) [3]. The basic idea of CF algorithm is to apply common experiences or similar preference to suggest what the user may be interested in, but not revealed yet. Its main process is to find the user that is similar to the target user or find the item similar to the predicted item. The advantages of having a collaborative recommendation system are people can discover new interest about the product and there are no special requirements for the recommended resources. However, CF algorithm still has some defect such as cold start problem [4]. This paper we present movie recommendation system using social network analysis and *k*-nearest neighbor

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 606–611, 2020.
https://doi.org/10.1007/978-981-13-9341-9_104

algorithm. We used the centrality of social network analysis i.e. betweenness centrality. our idea is clustering by using community detect based on edge betweenness centrality for the user and applied *k*-NN to finding the group for new user with their similar characteristic information such as gender, age, and occupation. The system will recommend movies in the group with similarity to target users. The purpose of this study was to develop techniques that can recommend the most appropriate movies to new users based on their personal characteristics. The main motivation of this paper is to develop a technique that recommends the movies according user's preference.

2 Related Works

Now a day, recommender systems have become increasingly popular and it has been used in various areas such as movies, music, news, book, products in general. Collaborative filtering (CF) is a method to make automatically predicts or filters information about a user's interests by compiling preferences or taste information from many users. To understand what CF is, we can think of a simple question, for example, if you want to watch a movie now, but you don't know which movies is good, the most people like to ask their friends which movie is good o and in general, we like to get suggestions from friends who have the same taste. This is the main idea of CF method [5].

2.1 k-Nearest Neighbor Algorithm

k-Nearest Neighbor or *k*-NN is an algorithm that used for classification of a categorical outcome or prediction of a numerical outcome. To classify or predict a new record, the method relies on finding similar record in the training data. These neighbor are then used to derive a classification or averaging for prediction. *k* in here represents the maximum number of neighbors to consider in counting the labels of the neighbors. Euclidean distance is also calculated to find the neighbors to consider. Details on this are given as follows: the technique of the *k*-NN algorithm is to find the numerical distance between each variable or attributes in the data. The distance computed by Euclidean distance as follows [6]:

$$\mathbf{dist}(\boldsymbol{p}, \boldsymbol{q}) = \sqrt{\sum_{i=1}^{n} (\boldsymbol{p}_i - \boldsymbol{q}_i)^2} \tag{1}$$

Which n is the number of dimensions (attributes), p and q denote two different users of the interest previously and p_i and q_i are, respectively, the attributes (components) or data objects p and q.

2.2 Social Network Analysis

A social network is a process in investigating social structures using networks and graph theory. It characterizes networks structure in terms of nodes like actor individual, people, or thing in the network. Centrality is representative indicator used in social network analysis. There are 3 kind of centrality including degrees' centrality,

betweenness centrality and closeness centrality. Betweenness centrality is the measure of the center of the graph based on the shortest path. The betweenness centrality for each vertex is the number of the shortest paths that pass through that vertex. Girven et al., presented community detect in social network and biological network based on edge betweenness centrality to avoid the shortcomings of the hierarchical clustering method, they find which edge in the network are most betweenness other pair of vertices by used betweenness centrality to edges and determine the edge betweenness of edge like the shortest path number between pair of vertices that runs long it [7]. The algorithm used for identifying communities is simply stated as follow:

(1) Calculate the betweenness centrality for all edges in the network.
(2) Remove the edge with the highest betweenness.
(3) Recalculate betweennesses centrality for all edges affected by the removal.
(4) Repeat from step 2 until no edges remain.

3 Movie Recommendation System Using Social Network Analysis and *k*-Nearest Neighbor Algorithm

The illustration of system movie recommendation system using social network analysis and *k*-nearest neighbor is shown in Fig. 1.

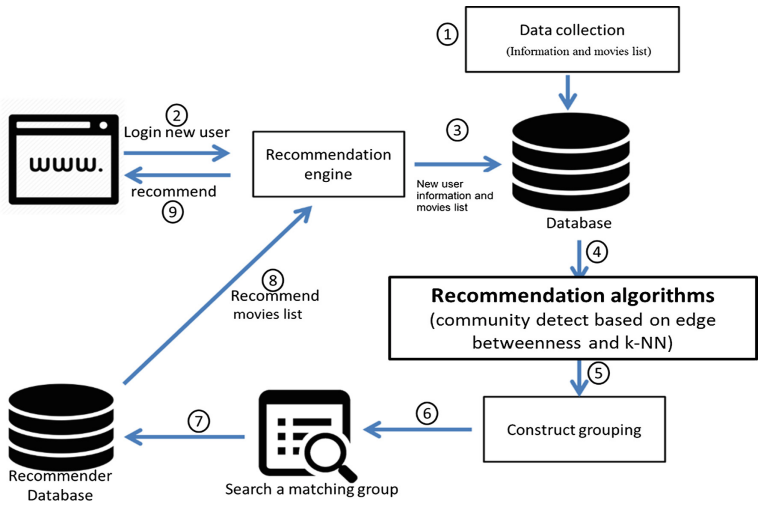


Fig. 1. The illustration of movie recommendation system using Social network analysis and *k*-NN

Number ① in Fig. 1, the system needs to collect user's information and list of movies into the database. Number ② in Fig. 1, when new users want to join to the system they need to login. Number ③ in Fig. 1, new user information needs to collect

in to database. Number ④ in Fig. 1, we find similar between from their personal characteristics, and then converting them to adjacency matrix. Number ⑤ in Fig. 1, in this step, the adjacency matrix is used for cluster users into several groups by applied community detection based on edge betweenness. Number ⑥ in Fig. 1, *k*NN algorithm helps find a group of new user by compute distance between new user and users in each group. The group that has the least distance value is selected for the new user. Number ⑦ in Fig. 1, the list movies that watched by users in the matching group are stored in the database. Number ⑧ in Fig. 1, the movies that watched by the users in group will sorting in order of popularity. Number ⑨ in Fig. 1, recommended top- watched movies in the group to the new user.

4 Experimental Analysis

4.1 Experiment Dataset

In our study, we used MovieLens dataset. The dataset was provided by GroupLens research group in the University of Minnesota, America [8]. It comprises data from three orders of scale and each dataset is involved about user's information, movies information user's ratings. The dataset are consists of 100,000 rating (the range of rating is 1 to 5) from 943 users on 1,682 movies; each user rated at least 20 movies and simple demographic information, such as gender, age and occupation is provided. In our experimental, we have divided the dataset into 2 parts: training and testing dataset, for training dataset is used 85% and testing is 15%.

4.2 Experimental Results

To evaluate the recommendation accuracy of the algorithm, the most commonly used measurement Mean Absolute Error (MAE) in CF recommender systems is used. MAE can measure the prediction accuracy by calculating the deviation between user prediction and real data. The smaller MAE is the higher accuracy of the recommendation. MAE is calculated as follows [9].

$$\mathrm{MAE} = \frac{1}{n}\sum_{\mathrm{i}=1}^{\mathrm{n}} |r_i - v_i| \tag{2}$$

Where n is the number of books read by the target users, r_i is the actual book of the user, v_i denote the predicted books of the recommendation system.

When performing our proposed method, we want to predict how accurate of our movie recommendation using social network analysis and *k*-NN is. We used MAE help to measure the satisfaction degree and estimate the accuracy of the recommender system. Usually, the smaller the MAE value, the higher the recommendation accuracy. The MAE results of our movie recommendation system using social network and *k*-nearest neighbor is shown in Fig. 2.

In Fig. 2 when we recommended 5 movie to the target user the MAE was 1.82517 and it was 1.93846, 2.34440, 2.48181, 2.613461 respectively. We can say that, the best number of recommend was 5 movies.

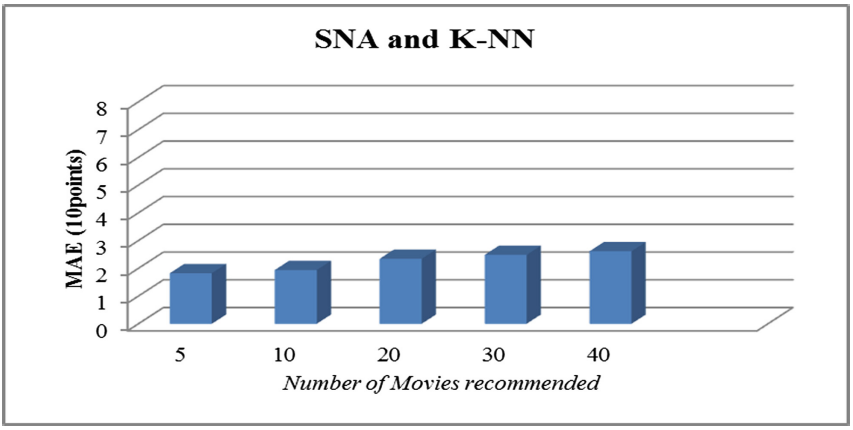


Fig. 2. The MAE result of movie recommender system using social network and *k*-nearest neighbor

For the purpose to prove our proposed algorithm, we compare the result of movie recommendation system and k-nearest neighbor with four experimental algorithms including movie recommendation system using social network analysis and collaborative, movie recommendation system using density-based on clustering, movie recommendation system using *k*-nearest neighbor and collaborative filtering, and we also compare with movie recommendation system using original collaborative filtering as shown in Fig. 3. the result compared shown our proposed algorithm very effective for movie recommendation and more accuracy than four algorithms that we mention previously.

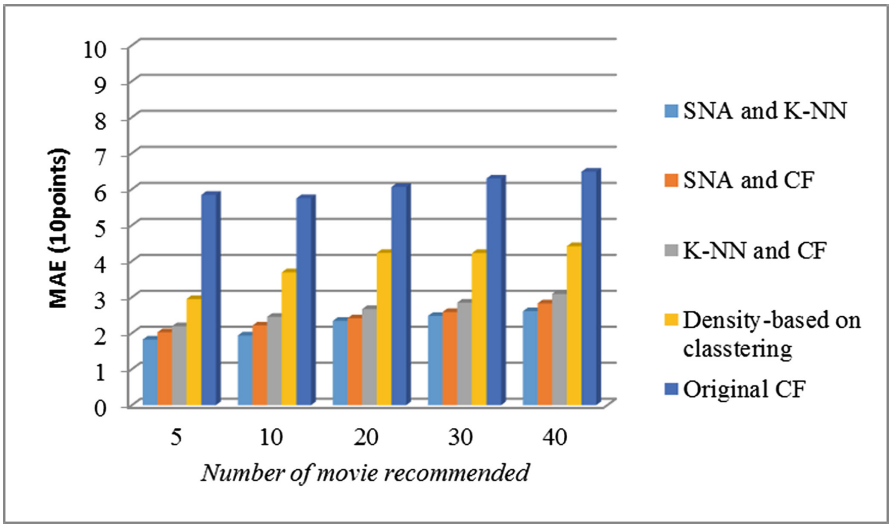


Fig. 3. Comparative result

5 Conclusion

In this paper, we have proposed an alternative approach for recommendation system using social network and *k*-nearest neighbor algorithm to solve the problem of collaborative filtering. We clustering the community or group to user based on edge betweenness centrality and *k*-NN helps to find users similar to the target user by using their personal information. The method that we proposed in this paper is very effective for movie recommendation system. Analyzing results shown that the best number of recommended is 5 movies which the MAE was 1.82517 from maximum 10 point. In addition, the method presented in this paper showed the best performance, followed by social network analysis and CF, followed *k*-NN and CF, Density-based clustering, and CF respectively.

Acknowledgements. This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under ITRC (Information Technology Research Center) support program (IITP-2018-2014-1-00720) supervised by IITP (Institute for Information & communications Technology Promotion) and the National Research Foundation of Korea (No. NRF-2017R1A2B1008421).

References

1. Park, D.H., Kim, H.K., Choi, I.Y., Kim, J.K.: A review and classification of recommender systems research. In: 2011 International Conference on Social Science and Humanity, IPDR vol. 5 (2011). IACSIT Press, Singapore © (2011)
2. Wang, J., Zhang, N.Y., Yin, J.: Collaborative filtering recommendation based on fuzzy clustering of user preference. In: 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2010) (2010)
3. Ye, F., Zhang, H.: A collaborative filtering recommendation based on users' interest and correlation of items. In: International Conference on Audio, Language and Image Processing (ICALIP), July 2016
4. Tiwari, S.K., Shivastava, S.K.: An approach for recommender system by combining collaborative filtering with user demographics and items genres. Int. J. Comput. Appl. **128**(13), 16–24 (2015). (0975-8887)
5. Yongchang, W., Ligu, Z.: Research on collaborative filtering recommendation algorithm based on mahout. In: 2016 4th International Conference on Applied Computing and Information Technology/3rd International Conference on Computational Science/Intelligence and Applied Informatics/1st International Conference on Big Data, Cloud Computing, Data Science & Engineering. IEEE (2016)
6. Aggarwal, C.C.: Recommender Systems. Springer, Berlin (2016)
7. Girvan, M., Newman, M.E.J.: Community structure in social and biological networks. PNAS **99**(12), 7821–7826 (2002)
8. Miller, B.N., Albert, I., Lam, S.K., et al.: MovieLens unplugged: experiences with an occasionally connected recommender system. In: Proceedings of the 8th International Conference on Intelligent User Interfaces, pp. 263–266. ACM (2003)
9. Parvatikan, S., Joshi, B.: Online book recommendation system by using collaborative filtering and association mining. In: IEEE International Conference on Computational Intelligence and Computing Research (ICCCI), December 2015

A Framework for Learning the Pricing Model of Sensing Tasks in Crowdphotographing

Fei Hao[1], Huijuan Guo[2], and Doo-Soon Park[3(✉)]

[1] School of Computer Science, Shaanxi Normal University, Xi'an, China
feehao@gmail.com
[2] Department of Computer Science, Taiyuan Normal University, Taiyuan, China
[3] Department of Computer Software Engineering,
Soonchunhyang University, Asan, Korea
parkds@sch.ac.kr

Abstract. Crowdphotographing, an emerging self-service mode over the mobile Internet, is to recruit several users to take the pictures via incentive mechanism. Importantly, it can be used for business inspection and information collection for enterprises. This paper mainly studies the rationality and optimization of task pricing for crowdphotographing. In this paper, different multivariate linear regression models are established to analyze the task pricing, performance, membership information and so forth. The multivariate linear regression equation is eventually obtained, which efficiently solves the problem of task pricing.

Keywords: Crowdphotographing · Pricing model · Multivariate linear regression

1 Introduction

Crowdsourcing as a promising services computing paradigm, is to make full use of human resources and massive human intelligence for various crowdsourcing services [6], such as data sensing [1], multimedia optimization [2], and image classification [3] and so forth. Particularly, Crowdphotographing, an emerging self-service mode over the mobile Internet, allows participants to take the necessary pictures freely for further analysis and services mining [7, 8]. Then, participants can obtain some award according to the incentive mechanism provided by the crowdphotographing platform.

The users can download the crowdphotographing APPs, and registers as a member of it, and then receives the photographic task from the APP, earning the money that the APP sets for the task. As the best case of the new self-service mode, "making money by

This research was supported by the National Natural Science Foundation of China (Grant No. 61702317), MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2018-2014-1-00720) supervised by the IITP (Institute for Information & communications Technology Promotion) and the National Research Foundation of Korea (No. 2017R1A2B1008421) and was also supported by the Fundamental Research Funds for the Central Universities, China (GK201703059).

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 612–617, 2020.
https://doi.org/10.1007/978-981-13-9341-9_105

taking pictures" has greatly saved the cost of investigation, ensured the authenticity of survey data and shortened the investigation cycle. Therefore, APP becomes the core of the platform operation medium, while task pricing is the core research issue. If task pricing within a reasonable range, then the number of visitors will fight for completing it; On the contrary, if the price is unreasonable, then these tasks will be ignored, resulting in the completion of the task is not high, and close to the failure of commodity inspection. Therefore, how to determine the task pricing model is becoming a major challenge.

Aiming to decide an efficient task pricing model for crowdphotograhing, this paper made the following major contributions: (1) Regarding to a real-world dataset, we use a the multivariate linear regression models to model the task pricing; (2) Based on these models, we further extract the factors set by establishing a scatter plot of task pricing and four factors including the longitude, latitude, the number of members and the creditworthiness of member; (3) The relevant simulations are conducted for evaluating the effectiveness of the proposed model.

The rest of this paper is organized as follows. Section 2 describes the experimental dataset about crowdphotographing. Then, the task pricing model is presented in Sect. 2. Section 3 concludes this paper.

2 A Framework for Learning the Pricing Model of Sensing Tasks in Crowdphotographing

In this section, a framework for learning the pricing model of sensing tasks in crowdphotographing is provided. It is composed of the following technical aspects: (1) Correlation between tasks and locations of members; (2) The best radius to distribution of members; (3) The best radius for packing. (4) The creditworthiness of members. Based these analysis, we devise a multivariate linear regression model based pricing mechanism.

First, the experimental datasets are described. The experimental crowdphotographing dataset which is available at the website of China Undergraduate Mathematical Contest in Modeling (CUMCM)[1]. This dataset contains 3 files that are completed tasks data, members' data, and newly launched tasks data of a new project.

2.1 Correlation Between Tasks and Locations of Members

The visualization software bdp is utilized to establish a relationship diagram between the task and the member. Through the relationship diagram, the relationship between the two locations is more clearly seen. It is roughly judged that the task pricing and completion are closely related to the position of the member (the more the number of members, the task The lower the price, the better the completion), that is, the task pricing is inversely proportional to the number of members, and the completion is directly proportional to the number of members; thus further thinking about the direct

[1] http://www.mcm.edu.cn.

relationship between the credibility of the members around them, the scheduled start time of the task, and the predetermined limit. (The higher the credibility of the member, the better the task completion; the more the member's reservation limit is, the worse the completion is). The task completion is directly proportional to the member's credibility and inversely proportional to the member's predetermined limit (Fig. 1).

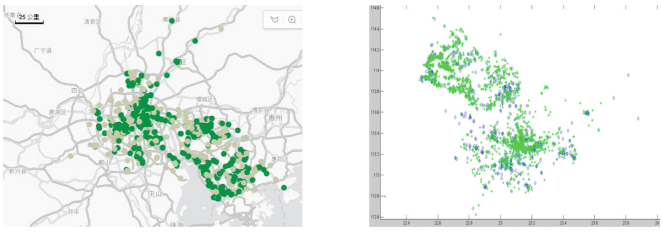


Fig. 1. The relationship diagram between tasks and members on the map and the coordinate system

As matter of fact, the more the number of members, the price of task is lower and the quality is better. On the contrary, the fulfillment is directly proportional to the number of members, and it is further thought that it is directly related to the credibility of the members around it, the scheduled time for the start of the task, and the prior limitation. It is directly proportional to the credibility of a member, and is inversely proportional to the prior limitation.

2.2 The Best Radius to Distribution of Members

In order to find the number of members in the vicinity of each task, the reputation value and the predetermined limit of the member who can complete the task are obtained, thereby calculating the probability of completion of the task, and finding the member information within a large range becomes a key issue. With the task as the center, the calculated optimal radius can ensure that each member can complete, and the radius is not too large, resulting in an excessive range, which makes the calculated task completion probability inaccurate. Here we use the optimization idea [4] and the stepwise approximation method to calculate the optimal radius centered on the task.

2.3 The Best Radius for Packing

Task packing are benefit to both members and crowdphotographing providers. For a user, he/she can obtain much reward. Let us assume one given task with the price 100 USD, if he/she wants to obtain 1000 USD, he/she will have to grab 10 tasks, and after packing, he will grab 10 tasks at a time, and the pay will be higher. If a business pays 100 USD for a task, 1000 USD for 10 tasks, and he can pay 800 USD for packaging, the cost could be reduced, but the charm of the task is increased.

How many tasks should be packaged properly? It depends on the difficulty of the task itself. In this research, the task packaging threshold is set to 10. If the threshold is less than 10, then we have to pack these tasks; If the threshold is greater than 10, then the nearby 10 tasks will be packaged.

2.4 The Creditworthiness of Members

The credibility of members' completion of tasks directly affects the quality of task completion. Naturally, the users with high credibility should be given priority in task competition, and there is no need to delegate tasks to an unreliable person. Therefore, it is necessary to establish a reasonable credit rating standard. Due to the limitation of the paper, the establishment of the rating standard is not the problem considered in this paper.

2.5 Multivariate Linear Regression Model Based Pricing Mechanism

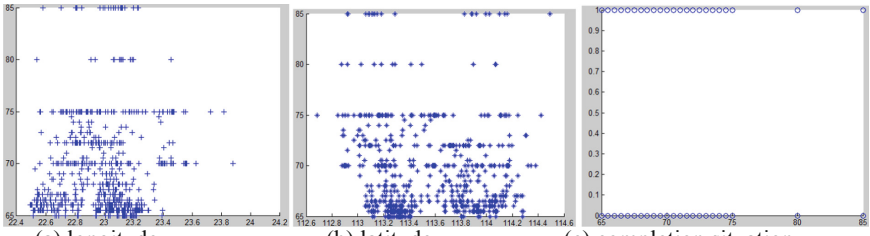
This section utilizes the multivariate linear regression model for capturing the correlation among the factors which affect the pricing model for crowdphotographing.

Assuming that the explanatory variable Y has a linear relationship with multiple explanatory variables $X_1, X_2, \ldots X_k$, it is a multivariate linear function of explanatory variables, called multivariate linear regression model. That is,

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots \beta_k X_k + \mu \tag{1}$$

where Y is the explanatory variable, X_j (j = 1, 2,…, k) is the K interpreted variable, and β_j (j = 0, 1, 2,…, k) refers to K + 1 unknown parameter, μ indicates the random error.

According to the first file in our experimental dataset, the scatter diagram of task pricing, longitude, latitude and task completion situation are shown in Fig. 2(a), (b), (c).



(a) longitude (b) latitude (c) completion situation

Fig. 2. The scatter diagrams between tasks pricing and longitude, latitude, and completion situation

It can be concluded from the above scatter diagrams that task pricing is not simply determined by latitude and longitude of tasks. Therefore, we not only consider the longitude and latitude of the task, but also the number of members around the task and the average reputation of members around the task, and then an efficient linear equation can be established as follows.

Step 1:
Dependent variable: task pricing **y**
Independent variables: longitude of task $\mathbf{x_1}$, dimension of task $\mathbf{x_2}$, number of members $\mathbf{x_3}$ and average reputation $\mathbf{x_4}$ of members around the task within a certain radius

Step 2: calculating the number of members around task *t* within the radius r.

Step 3: within the circle with radius r, we calculate the average reputation value of those members around the given task *t*.

Step 4: we build a multivarible linear regression equation

$$y = k_0 + k_1x_1 + k_2x_2 + k_3x_3 + k_4x_4$$

Based on the Least Square Method (LSM) [5], we estimate the values of k_0, k_1, k_2, k_3 and k_4 with the help of MATLAB software. The results are as follows:

(1) If only the longitude and latitude of the task are considered for the task pricing, the multivariate linear regression equation can be established with the task pricing as the dependent variable and the task longitude and latitude as the independent variable. The multivariable linear regression equation is

$$y = k_0 + k_1x_1 + k_2x_2$$

Similarly, we estimate the parameters with MATLAB software and get the multivariable linear regression equation as follows.

$$\mathrm{y} = -9.4375 + 2.476796\mathrm{x}_1 + 0.278544\mathrm{x}_2$$

(2) If we consider the longitude, latitude of the task, the number of members within a certain radius, and average value of reputation for the task pricing, we similarly estimate the parameters with MATLAB software and get the multivariable linear regression equation as follows.

$$\mathrm{y} = -215.329 + 4.769025\mathrm{x}_1 + 1.544462\mathrm{x}_2 - 0.21145\mathrm{x}_3 - 0.04794\mathrm{x}_4$$

The model is applied to predict the unfinished tasks. The results show that the task completion rate increases to 94%, which is more than 30% points higher than before. The results show that the improved model can reasonably price the task when the factors are more sufficient, and the multivariable linear regression equation is more accurate.

3 Conclusions

This paper aims to present an efficient framework of pricing model for efficient crowdphotographing, different multivariate linear regression models are established to analyze the task pricing, performance, membership information and so forth. The multivariate linear regression equation is eventually obtained, which efficiently solves the problem of task pricing. The experimental results show that the task completion rate could be increased to 94%, which is more than 30% points higher than before.

References

1. Zhang, X., Wu, Y., Huang, L., et al.: Expertise-aware truth analysis and task allocation in mobile crowdsourcing. In: IEEE International Conference on Distributed Computing Systems, pp. 922–932. IEEE (2017)
2. Wu, T., Dou, W., Ni, Q., Yu, S., Chen, G.: Mobile live video streaming optimization via crowdsourcing brokerage. IEEE Trans. Multimedia **19**(10), 2267–2281 (2017)
3. Tran-Thanh, L., Venanzi, M., Rogers, A., et al.: Efficient budget allocation with accuracy guarantees for crowdsourcing classification tasks. In: Twelfth International Conference on Autonomous Agents and Multi-Agent Systems, pp. 901–908 (2013)
4. Tang, H., Sun, Z.C., Chew, K.W.R., et al.: A 5.8 nW 9.1-ENOB 1-kS/s local asynchronous successive approximation register ADC for implantable medical device. IEEE Trans. Very Large Scale Integr. Syst. **22**(10), 2221–2225 (2014)
5. Dong, L., Zhou, J., Tang, Y.Y.: Effective and fast estimation for image sensor noise via constrained weighted least squares. IEEE Trans. Image Process. **27**(6), 2715–2730 (2018)
6. Tong, Y., Chen, L., Zhou, Z., et al.: SLADE: a smart large-scale task decomposer in crowdsourcing. IEEE Trans. Knowl. Data Eng. **30**(8), 1588–1601 (2018)
7. Hao, F., Jiao, M., Min, G., et al.: Launching an efficient participatory sensing campaign: a smart mobile device-based approach. ACM Trans. Multimedia Comput. Commun. Appl. **12** (1s), 18 (2015)
8. Hao, F., Jiao, M., Min, G., et al.: A trajectory-based recruitment strategy of social sensors for participatory sensing. IEEE Commun. Mag. **52**(12), 41–47 (2014)

Tensor Decomposition Based Electrical Data Recovery

Shiming He[1(✉)], Zhuozhou Li[1], Jin Wang[1], Kun Xie[2], and Dafang Zhang[2]

[1] School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha 410114, China
{smhe_cs, jinwang}@csust.edu.cn

[2] College of Computer Science and Electronics Engineering, Hunan University, Changsha 410082, China

Abstract. As the development of smart grid and energy internet, the amount of transmitted data in real time significantly increase. Due to the mismatch with communication networks that were not designed to carry high-speed and real time data, data losses and data quality degradation may happen constantly. For this problem, according to the strong spatial and temporal correlation and periodicity of electricity data which is generated by human's actions and feelings, we treat the electricity data as a tensor where the three dimensional are user, weeks, days. We divide the electricity data tensor into the sum of multiple rank-1 tensors and use the known data to approximate the electricity data tensor and recover the lost electrical data. Based on the real electricity data, we analyze the sparseness of the electricity data tensor and perform the CP decomposition-based method on the real data. The experimental results verify the recovery efficiency of the proposed scheme.

Keywords: Electrical data recovery · Tensor factorization · Sparse

1 Introduction

As the development of smart grid and energy internet [1], the amount of transmitted data in real time significantly increase. Due to the mismatch with communication networks that were not designed to carry high-speed and real time data, data losses and data quality degradation may happen constantly.

For this problem, according to the strong spatial and temporal correlation of electricity data which is generated by human's actions and feelings, some works treat the electricity data as a low-rank matrix and use Matrix factorization (MF) [2] to recovery lost data. And some work takes the weather information as aid to recover electrical data via collective matrix factorization [3].

However, the weather information is quiet different for different locations, which can only be used to recover the electrical data of one location [3]. They ignore that the periodicity of human's actions and feelings, such the weekday and weekend.

Inspired by tensor factorization, we treat the electricity data as a tensor where the three dimensional are days, user, weeks. we divide the electricity data tensor into the

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 618–624, 2020.
https://doi.org/10.1007/978-981-13-9341-9_106

sum of multiple rank-1 tensors and use the known data to approximate the electricity data tensor and recover the lost electrical data. Based on the real electricity data, we analyze the sparseness of the electricity data tensor and perform the CP [4] decomposition-based method on the real data. The experimental results verify the recovery efficiency of the proposed scheme.

2 System Model

Generally, the value of the smart meter is the the cumulative power consumption of user on each day. The minus of two consecutive values is the power consumption on one day. We take the power consumption on each day as the electricity data. The electricity data is generated by human's actions and feelings, which has a strong spatial and temporal correlation. At the same time the human's actions and feelings has periodicity, such the weekday and weekend. Therefore, we treat the electrical data as a three-order or three-way tensor $\mathcal{X} \in \mathbb{R}^{W \times N \times H}$, instead of a matrix. In the electrical tensor, there are N uses and H weeks to consider with each week having W days (W = 7). The electrical data tensor contains the data within a W**H* times measurement for *N* users. An element x_{ijk} represents the power consumption of user *j* on the *i*th day of the *k*th week, as shown in Fig. 1.

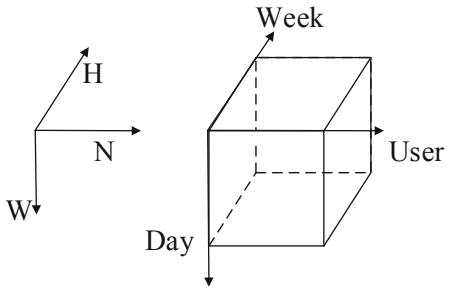


Fig. 1. The three-order tensor of electrical data. The first-way is the days of a week, the second-way is the users, and the third-way is the number of weeks.

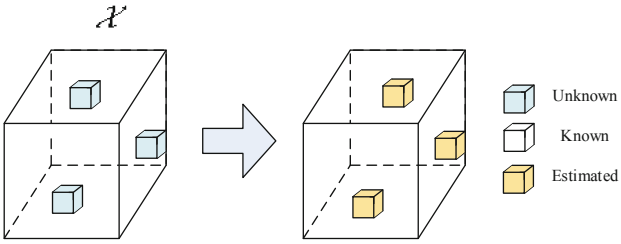


Fig. 2. The recovery of electrical data.

The electrical tensor has many lost elements. The subset Ω of tensor is the known set, where the elements x_{ijk}, $(i,j,k) \in \Omega$ are known. As shown in Fig. 2, the recovery task is to estimate the lost or unknown element in the tensor by the spatial and temporal correlation and periodicity of the data in order to minimize the recovery error, which is usually defined as squared error $(x_{ijk} - \hat{x}_{ijk})^2$, where x_{ijk} is the real value and $\hat{x}_{ijk}$ is the estimated value.

3 CP-Decomposition Based Electrical Data Recovery

3.1 CP-Decomposition

CP-decomposition [4] is also called as CANDECOMP/PARAFAC model. Inspired of matrix fraction [2] ($x_{ij} = \sum_{r=1}^{R} a_{ir} b_{ir}$), in CP-decomposition a element of a three-way tensor can be calculated by the sum of a finite number of multiply of three scalars, as shown in Fig. 3.

$$x_{ijk} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr} \tag{1}$$

Where R > 0. Then the three-way tensor can be expressed as the sum of R rank-1 tensors which can be calculated by the outer product of three vectors.

$$\mathcal{X} = \sum_{r=1}^{R} \boldsymbol{a}_r \circ \boldsymbol{b}_r \circ \boldsymbol{c}_r \tag{2}$$

where $\boldsymbol{a}_r = [a_{1r}, \ldots, a_{Wr}]^T \in \mathbb{R}^{W \times 1}, \boldsymbol{b}_r = [b_{1r}, \ldots, b_{Nr}]^T \in \mathbb{R}^{N \times 1}, \boldsymbol{c}_r = [c_{1r}, \ldots, c_{Hr}]^T \in \mathbb{R}^{H \times 1}$ are the factor vectors, $\circ$ is the out-product of vectors. Further, we can get the tenor's factor matrices.

$$\mathbf{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_R] \in \mathbb{R}^{W \times R}, \mathbf{B} = [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_R] \in \mathbb{R}^{N \times R}, \mathbf{C} = [\boldsymbol{c}_1, \ldots, \boldsymbol{c}_R] \in \mathbb{R}^{H \times R} \tag{3}$$

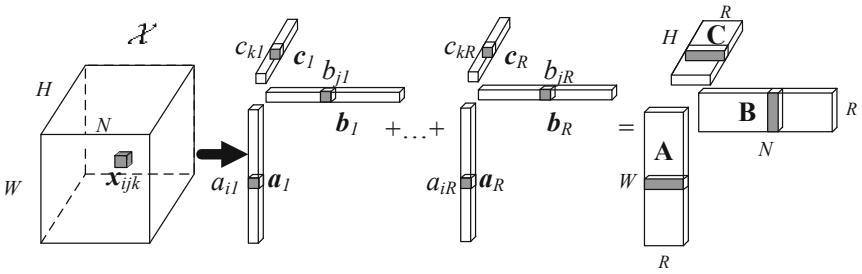


Fig. 3. The CP-decomposition.

According to the kolda horizontal unfolding, the CP-decomposition can be also expressed as following.

$$\begin{aligned}\mathbf{X}^{(W\times NH)} &= \mathbf{A}(\mathbf{C}\odot\mathbf{B})^{\mathrm{T}},\\ \mathbf{X}^{(N\times WH)} &= \mathbf{B}(\mathbf{C}\odot\mathbf{A})^{\mathrm{T}},\\ \mathbf{X}^{(H\times WN)} &= \mathbf{C}(\mathbf{B}\odot\mathbf{A})^{\mathrm{T}},\end{aligned} \tag{4}$$

where $\odot$ is the Khatri-Rao product of matrices. $\mathbf{X}^{(W\times NH)}, \mathbf{X}^{(N\times WH)}, \mathbf{X}^{(H\times WN)}$ are the horizontal unfolding of tensor with forntal, horizontal, and lateral slices. The following is a deducing process.

$$\begin{aligned}\mathbf{X}^{(W\times NH)} = [\mathbf{X}_{::1},\ldots,\mathbf{X}_{::H}] &= [\boldsymbol{a}_1\boldsymbol{b}_1^{\mathrm{T}}c_{11}+\ldots\boldsymbol{a}_R\boldsymbol{b}_R^{\mathrm{T}}c_{1R},\ldots,\boldsymbol{a}_1\boldsymbol{b}_1^{\mathrm{T}}c_{H1}+\ldots\boldsymbol{a}_R\boldsymbol{b}_R^{\mathrm{T}}c_{HR}]\\ &= [\boldsymbol{a}_1,\ldots,\boldsymbol{a}_R]\begin{bmatrix} c_{11}\boldsymbol{b}_1^{\mathrm{T}}\ldots c_{H1}\boldsymbol{b}_1^{\mathrm{T}}\\ \ldots\\ c_{1R}\boldsymbol{b}_R^{\mathrm{T}}\ldots c_{HR}\boldsymbol{b}_R^{\mathrm{T}}\end{bmatrix} = \mathbf{A}(\mathbf{C}\odot\mathbf{B})^{\mathrm{T}}\end{aligned} \tag{5}$$

where $\mathbf{X}^{(W\times NH)}$ is constructed by H forntal slices, $\mathbf{X}_{::h}$ is hth forntal slice. Later, we need to use the unfolding to solve the CP recovery problem.

3.2 CP-Decomposition for Recovery

The object of the CP-decomposition for recovery is to minimize the error sum of the known elements, which is as follows.

$$\min_{\mathbf{A},\mathbf{B},\mathbf{C}}\frac{1}{2}\sum_{(i,j,k)\in\Omega}(x_{ijk}-\hat{x}_{ijk})^2 = \min_{\mathbf{A},\mathbf{B},\mathbf{C}}\frac{1}{2}||(\mathcal{X}-\sum_{r=1}^{R}\boldsymbol{a}_r\circ\boldsymbol{b}_r\circ\boldsymbol{c}_r)_{\Omega}||_{\mathrm{F}}^2 \tag{6}$$

where $||\bullet||_{\mathrm{F}}$ represents Frobenius norm. The object function is not joint convex of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$. Therefore, we can use the alternating least squares (ALS) to solve it. According to the kolda unfolding, problem (6) can be divided into three sub-optimal problem.

$$\begin{aligned}\mathbf{A} &= \arg\min_{\mathbf{A}}||\mathbf{X}^{(W\times NH)}-\mathbf{A}(\mathbf{C}\odot\mathbf{B})^{\mathrm{T}}||_{\mathrm{F}}^2,\\ \mathbf{B} &= \arg\min_{\mathbf{B}}||\mathbf{X}^{(N\times WH)}-\mathbf{B}(\mathbf{C}\odot\mathbf{A})^{\mathrm{T}}||_{\mathrm{F}}^2,\\ \mathbf{C} &= \arg\min_{\mathbf{C}}||\mathbf{X}^{(H\times WN)}-\mathbf{C}(\mathbf{B}\odot\mathbf{A})^{\mathrm{T}}||_{\mathrm{F}}^2,\end{aligned} \tag{7}$$

To let the derivation be zero, we can obtain the ALS solutions of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$.

$$\begin{aligned}\mathbf{A} &= \mathbf{X}^{(W\times NH)}((\mathbf{C}\odot\mathbf{B})^{\mathrm{T}})^{+}\\ \mathbf{B} &= \mathbf{X}^{(N\times WH)}((\mathbf{C}\odot\mathbf{A})^{\mathrm{T}})^{+}\\ \mathbf{C} &= \mathbf{X}^{(H\times WN)}((\mathbf{B}\odot\mathbf{A})^{\mathrm{T}})^{+}\end{aligned} \tag{8}$$

where $^{+}$ presents the pseudo inverse matrix. According to the Khatri-Rao product's Moore-Penrose matrix character in Eq. (9), we can get the solutions as Eq. (10).

$$(\mathbf{C} \odot \mathbf{B})^{+} = (\mathbf{C}^{\mathrm{T}}\mathbf{C} * \mathbf{B}^{\mathrm{T}}\mathbf{B})^{+}(\mathbf{C} \odot \mathbf{B})^{\mathrm{T}} \tag{9}$$

$$\begin{aligned} \mathbf{A} &\leftarrow \mathbf{X}^{(W\times NH)}(\mathbf{C} \odot \mathbf{B})(\mathbf{C}^{\mathrm{T}}\mathbf{C} * \mathbf{B}^{\mathrm{T}}\mathbf{B})^{+} \\ \mathbf{B} &\leftarrow \mathbf{X}^{(N\times WH)}(\mathbf{C} \odot \mathbf{A})(\mathbf{C}^{\mathrm{T}}\mathbf{C} * \mathbf{A}^{\mathrm{T}}\mathbf{A})^{+} \\ \mathbf{C} &\leftarrow \mathbf{X}^{(H\times WN)}(\mathbf{B} \odot \mathbf{A})(\mathbf{B}^{\mathrm{T}}\mathbf{B} * \mathbf{A}^{\mathrm{T}}\mathbf{A})^{+} \end{aligned} \tag{10}$$

All above, the ALS Algorithm of CP-decomposition for recovery is shown as Algorithm 1.

Algorithm 1 ALS Algorithm of CP-decomposition for recovery

Input: $\mathcal{X} \in \mathbb{R}^{\mathrm{W}\times N\times H}$, $\mathbf{R}$, Error threshold $\mathcal{E}$

Output: $\mathbf{A} \in \mathbb{R}^{\mathrm{W}\times R}, \mathbf{B} \in \mathbb{R}^{\mathrm{W}\times R}, \mathbf{C} \in \mathbb{R}^{\mathrm{W}\times R}$

1.Random initialization $\mathbf{A}_0 \in \mathbb{R}^{\mathrm{W}\times R}, \mathbf{B}_0 \in \mathbb{R}^{\mathrm{W}\times R}, \mathbf{C}_0 \in \mathbb{R}^{\mathrm{W}\times R}$

2. Calculate the kolda unfolding of tensor $\mathbf{X}^{(W\times NH)}, \mathbf{X}^{(N\times WH)}, \mathbf{X}^{(H\times WN)}$

3. While $k = 1,2,...$ && $\| \mathbf{X}^{(W\times NH)} - \mathbf{A}_{k+1}(\mathbf{C}_{k+1} \odot \mathbf{B}_{k+1})\| < \varepsilon$

$\mathbf{A}_{k+1} \leftarrow \mathbf{X}^{(W\times NH)}(\mathbf{C}_k \odot \mathbf{B}_k)(\mathbf{C}_k^{\mathrm{T}}\mathbf{C}_k * \mathbf{B}_k^{\mathrm{T}}\mathbf{B}_k)^{+}$

4. $\mathbf{B}_{k+1} \leftarrow \mathbf{X}^{(N\times WH)}(\mathbf{C}_k \odot \mathbf{A}_{k+1})(\mathbf{C}_k^{\mathrm{T}}\mathbf{C}_k * \mathbf{A}_{k+1}^{\mathrm{T}}\mathbf{A}_{k+1})^{+}$

$\mathbf{C}_{k+1} \leftarrow \mathbf{X}^{(H\times WN)}(\mathbf{B}_{k+1} \odot \mathbf{A}_{k+1})(\mathbf{B}_{k+1}^{\mathrm{T}}\mathbf{B}_{k+1} * \mathbf{A}_{k+1}^{\mathrm{T}}\mathbf{A}_{k+1})^{+}$

5.Return $\mathbf{A}_k, \mathbf{B}_k, \mathbf{C}_k$

4 Simulation

The real electrical data comes from Lanzhou power system company with 160 users in Jiuquan of the Lanzhou province from August 1, 2016 to August 31, 2017. Except for the lost data, we can get the available real data of 160 users in 385 days which are all known. The real data is treated as a tensor $\mathcal{X} \in \mathbb{R}^{7\times 160\times 55}$. There are 160 uses and 55 weeks with each week having 7 days.

The root mean squared error (RMSE) is used to evaluate the recovery accuracy, which is defined as:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{(i,j,k)\notin\overline{\Omega}} (x_{ijk} - \hat{x}_{ijk})^2}{|\overline{\Omega}|}} \tag{14}$$

where $\overline{\Omega}$ is set of the entries on which the values are unknown, $|\overline{\Omega}|$ is the number of unknown entries. If the RMSE is smaller, the recovery accuracy will be higher. We compare our scheme with the matrix factorization based recovery [3] on different sample ratios. The sample ratio is the ratio of the number of known elements to the

number of all elements in the electrical data. The higher the ratio, the more known elements, the more information we know, and the fewer elements we need to recover. We set the sampling ratios from 85% to 97.5%, increasing at 2.5% intervals.

Figure 4 shows the RMSE of CP and MF with different sample ratios. With the all sample ratio, the recovery accuracy of CP is better than that of MF. The reason is that the CP-decomposition uses more periodicity information, while the MF uses only spatial and temporal correlation. And as the sample ratio increases, the RMSE decreases because the more information is known, the more potential relationships will be provided to help improve recovery accuracy.

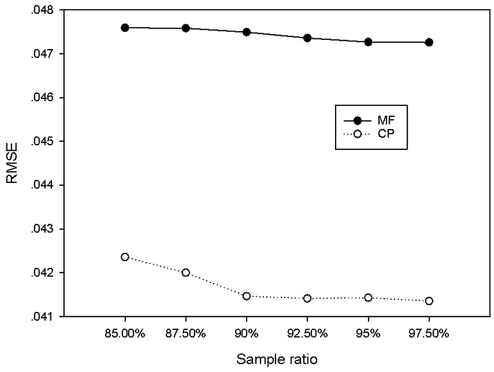


Fig. 4. The RMSE with different sample ratios

5 Conclusion

According to the strong spatial, temporal correlation and periodicity of electricity data, we treat them as a tensor where the dimensional are days, user, weeks. We analyze the sparseness of the electricity data tensor and perform the CP decomposition-based method on the real data. The experimental results on real data verify the recovery accuracy efficiency of the proposed scheme.

Acknowledgments. This work was supported by National Natural Science Foundation of China (Nos. 61802030, 61572184, 61502054), the Science and Technology Projects of Hunan Province (No. 2016JC2075), the Research Foundation of Education Bureau of Hunan Province, China (Nos. 16C0047, 16B085).

References

1. Tsoukalas, L.H., Gao, R.: From smart grids to an energy internet: assumptions, architectures and requirements. In: Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies, DRPT 2008, pp. 94–98. IEEE, New York (2008)

2. Tikk, D.: Investigation of various matrix factorization methods for large recommender systems. In: IEEE International Conference on Data Mining Workshops, pp. 1–6. IEEE, New York (2008)
3. Han, X., Dang, Q., Zhang, H., et al.: Electrical data recovery with weather information via collective matrix factorization. In: 2018 International Conference on Applications and Techniques in Cyber Intelligence (ATCI), pp. 1–10. Springer, Heidelberg (2018)
4. Kolda, T.G., Bader, B.W.: Tensor decomposition and applications. SIAM Rev. **51**(3), 455–500 (2009)

Compressing Deep Neural Network

Shiming He[1(✉)], Zhuozhou Li[1], Jin Wang[1], Kun Xie[2], and Dafang Zhang[2]

[1] School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha 410114, China
{smhe_cs, jinwang}@csust.edu.cn
[2] College of Computer Science and Electronics Engineering, Hunan University, Changsha 410082, China

Abstract. Deep learning is the most useful tool for may applications, such as image recognize, nature language processing. But huge computation power and millions of parameters are needed in large models which may can't be supported and stored. For this problem, some works tried to compress the dense weight matrices with sparse representations technologies, such as matrix decomposition and tensor decomposition. But it is still unknown which is the largest compress ratio. Therefore, in this paper, we analyse the relationship between the shape of tensor and the number of parameters, formulate the problem of minimizing the number of parameters, and solve it to find the best compress ratio. We compare the compressed ration on three data sets.

Keywords: Deep neural network · Parameters compressing · Matrix decomposition · Tensor decomposition

1 Introduction

Deep learning [1] is the most useful tool for may applications, such as image recognize, nature language processing. But huge computation power and millions of parameters are needed in large models which may can't be supported and stored. For this problem, some works tried to compress the dense weight matrices with sparse representations technologies. Denil et al. [2] employed low-rank matrix decomposition to replace the original weight matrix to compress the number of weight parameters. Further, Chien et al. [3] used the tucker decomposition to replace the affline transformation in an neural network. Novikov et al. [4] used Tensor Train decomposition format to represent the weight matrices in the fully connected layer.

Matrix decomposition and tensor decomposition are the general methods for compressing weight parameters. But it is still unknown which is the largest compress ratio. Therefore, in this paper, we analyse the relationship between the shape of tensor and the number of parameters, formulate the problem of minimizing the number of parameters, and solve it to find the best compress ratio. We compare the compressed ration on three data sets.

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 625–631, 2020.
https://doi.org/10.1007/978-981-13-9341-9_107

2 Tensor Decomposition Based Deep Neural Network

2.1 Unfolding Based DNN

Figure 1 is the structure of full connection layer with the transformations of $h = g(vW)$, where v is an input vector with I_1-dimensional input, h is an output vector with J_1-dimensional, and W is an $I_1 \times J_1$ matrix of parameters. And the number of parameters on this layer is $I_1 J_1$.

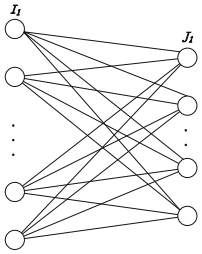


Fig. 1. The full connection layer of NN.

For matrix or tensor input, at first the matrix or tensor need to be unfolded into vector and then fed into the full connection layer. For a $I_1 \times I_2$ matrix input and $J_1 \times J_2$ matrix output, it is unfolded into an input vector with $I_1 I_2$-dimensional and an output vector with $J_1 J_2$-dimensional. The weight matrix is an $I_1 I_2 \times J_1 J_2$ matrix. And the number of parameters on this layer is $I_1 I_2 J_1 J_2$. Similarly, the three-way and N-way tensor input and output, the number of parameters are $I_1 I_2 I_3 J_1 J_2 J_3$ and $\prod_{n=1}^{N} I_n J_n$, respectively, after unfolding (Fig. 2).

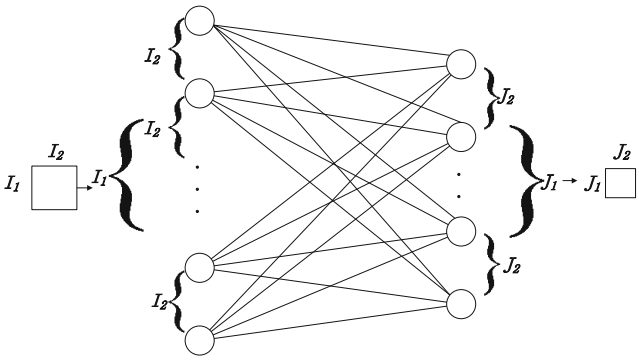


Fig. 2. The full connection layer on matrix input and output by unfolding.

2.2 Tensor Decomposition Based DNN

For a three-way tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, the tucker tensor decomposition [4] is to divide the tensor into the producer of a core-tensor and three factor matrices.

$$\mathcal{X} = \mathcal{A} \times_1 \mathrm{U}^{(1)} \times_2 \mathrm{U}^{(2)} \times_3 \mathrm{U}^{(3)} \tag{1}$$

where $\mathcal{A} \in \mathbb{R}^{J_1 \times J_2 \times J_3}$, $\mathrm{U}^{(1)} \in \mathbb{R}^{J_1 \times I_1}$, $\mathrm{U}^{(2)} \in \mathbb{R}^{J_2 \times I_2}$, $\mathrm{U}^{(3)} \in \mathbb{R}^{J_3 \times I_3}$ are the core-tensor and three factor matrices, $\times_i$ is the mode i produce (Fig. 3).

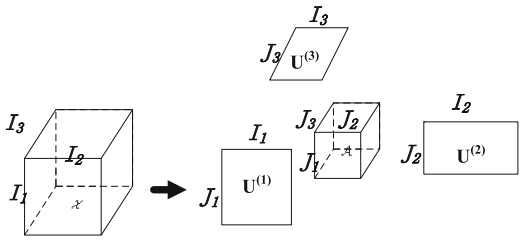


Fig. 3. The tucker tensor decomposition on three-way tensor.

According to the inverse of tucker tensor decomposition, we can take the tensor as input and output of layer directly [2], as shown in Fig. 4, where $^+$ presents the pseudo inverse matrix. There are $J_1J_2J_3$ weight tensors. All the weight tensors are the rank-1 tensor and they share the vectors which are the row of the pseudo inverse matrices from the tucker tensor decomposition.

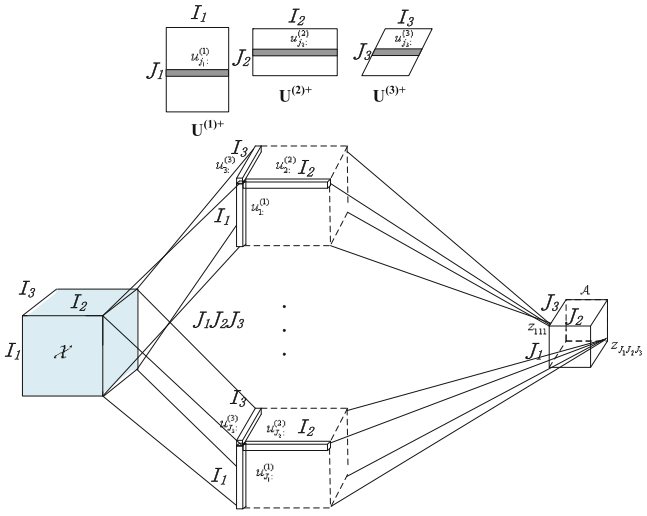


Fig. 4. The three-way tensor layer with tucker tensor decomposition.

The number of parameters of two-way (matrix) and three way tensor are $I_1J_1 + I_2J_2$, $I_1J_1 + I_2J_2 + I_3J_3$, respectively. For the n-way tensor, the number of parameters is $\sum_{n=1}^{N} I_nJ_n$, which is compressed comparing with unfolding in 2.1. The compressed ratio is $\prod_{n=1}^{N} I_nJ_n / \sum_{n=1}^{N} I_nJ_n$.

3 Tensor Reshaping

For the same dimensional output tensor, the different shape of input tensor affects the number of parameters and the compressed ratio.

For the three-way tensor, when the number of input elements is fixed which can be treated as the volume of the tensor, the number of weight parameters is the weight sum of length, width and height affect. For the n-way tensor, the problem of minimizing the number of weight parameters can be formulated as following.

$$\begin{aligned} &\min \sum_{n=1}^{N} I_n' J_n \\ &s.t. \prod_{n=1}^{N} I_n' = \prod_{n=1}^{N} I_n \end{aligned} \tag{2}$$

Where I_n and J_n is the n-th dimensional of input and output tensor, I_n' is the n-th dimensional of reshaped input tensor. To solve the problem (2), we can get the best links' length of reshaped input tensor.

$$I_i' = \left(\prod_{n=1}^{N} I_n \prod_{n=1, n\neq i}^{N} J_n \frac{1}{J_i^{N-1}} \right)^{1/N} \tag{3}$$

If all the links' length of output tensor J_n are equal, the number of weight parameters is dependent on the sum of the links' length of input tensor. Intuitively, the sum of the links' length of the super-cube is minimum, that is, $I_i' = \left(\prod_{n=1}^{N} I_n \right)^{1/N}$.

The compressing ratio can be reduced to

$$\begin{aligned}
&\prod_{n=1}^{N} I_n J_n / \sum_{n=1}^{N} I_n' J_n \\
&= \prod_{n=1}^{N} I_n J_n / (\sum_{i=1}^{N} \left(\prod_{n=1}^{N} I_n \prod_{n=1, n\neq i}^{N} J_n \frac{1}{J_i^{N-1}} \right)^{1/N} J_i) \\
&= \prod_{n=1}^{N} I_n J_n / \sum_{i=1}^{N} \left(\prod_{n=1}^{N} I_n \prod_{n=1, n\neq i}^{N} J_n \frac{J_i^N}{J_i^{N-1}} \right)^{1/N} \\
&= \prod_{n=1}^{N} I_n J_n / \sum_{i=1}^{N} \left(\prod_{n=1}^{N} I_n J_n \right)^{1/N} = \left(\prod_{n=1}^{N} I_n J_n \right)^{1-1/N} / N
\end{aligned} \tag{4}$$

4 Simulation

We use three data sets for evaluation. The Mixed National Institute of Standards and Technology (MNIST) database [5] consist of gray-scale images of size 28 × 28. The street view house number (SVHN) data set [6] is a real-world image data set, here the image size is 32 × 32 × 3. The TwoHandManip Gesture database [7] from video consist of images with spatial, temporal, and color information in a size of 8 × 12 × 10 × 3. The metric is the compressing ratio. Table 1 is the input size of the data sets. Table 2 is the output size of data sets.

Table 1. The input size of data sets.

	MNIST	SVHN	TwoHandManip gesture
Input size	28 × 28	32 × 32 × 3	8 × 12 × 10 × 3
Unfolding based DNN	784 × 1	3072 × 1	2880 × 1

Table 2. The reshaping result with different output size and data sets.

		Output size	Reshaping	
			Optimal value	Actual value
1	MNIST	20 × 20	28 × 28	28 × 28
2	SVHN	10 × 10 × 3	9.73 × 9.73 × 32.43	10 × 10 × 33
3	SVHN	20 × 20 × 3	7.72 × 7.72 × 51.49	8 × 8 × 51
4	TwoHandManip Gesture	4 × 6 × 2 × 3	6.34 × 4.22 × 12.68 × 8.45	6 × 4 × 13 × 9
5	TwoHandManip Gesture	4 × 6 × 5 × 1	6.06 × 4.04 × 4.84 × 24.24	6 × 4 × 5 × 24
6	TwoHandManip Gesture	4 × 6 × 5 × 3	7.97 × 5.31 × 6.38 × 10.63	8 × 5 × 7 × 11

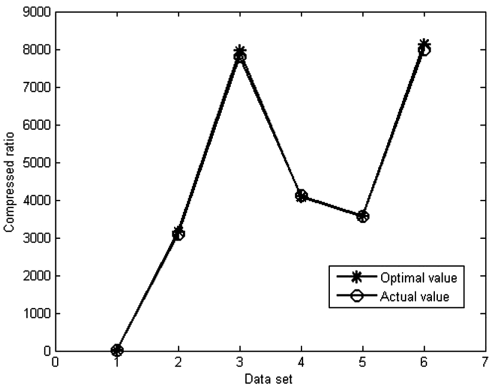


Fig. 5. The compressed ratio of different data sets.

Due to the optimal value from Eq. (3) may not be an integer, we should get the ceiling or floor of optimal value as the actual value. Therefore, the optimal value and actual value with different output size and data sets are shown in Table 2. We have two compressed ratios of optimal value and actual value in Fig. 5. There is a large compress ratio even with actual value.

5 Conclusion

According to the relationship between the shape of tensor and the number of parameters, we calculate the best compress ratio and reshape the input tensor for the tensor decomposition based DNN. We compare the compressed ration on three data sets. The future work is to verify the performance after reshaping.

Acknowledgments. This work was supported by National Natural Science Foundation of China (Nos. 61802030, 61572184, 61502054), the Science and Technology Projects of Hunan Province (No. 2016JC2075), the Research Foundation of Education Bureau of Hunan Province, China (Nos. 16C0047, 16B085).

References

1. Denil, B., Shakibi, L., Dinh, N., de Freitas et al., Predicting parameters in deep learning. In: Advances in Neural Information Processing Systems, pp. 2148–2156. IEEE Press, New York (2013)
2. Chien, J.T., Bao, Y.T.: Tensor-factorized neural networks. IEEE Trans. Neural Netw. Learn. Syst. **29**(5), 1998–2011 (2018)
3. Tjandra, A., Sakti, S., Nakamura, S., Compressing recurrent neural network with tensor train. In: 2017 International Joint Conference on in Neural Networks (IJCNN), pp. 4451–4458. IEEE Press, New York (2017)
4. Lathauwer, L.D., Moor, B.D., Vandewalle, J.: A multilinear singular value decomposition. SIAM J. Matrix Anal. Appl **21**(4), 1253–1278 (2000)

5. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998)
6. Netzer, Y., Wang, T., Coates, A., Bissacco, A., Wu, B., Ng, A.Y.: Reading digits in natural images with unsupervised feature learning. In: Proceedings of NIPS Workshop Deep Learning and Unsupervised Feature Learning, p. 5. IEEE Press. New York (2011)
7. Liao, C.P., Chien J.T.: Graphical modeling of conditional random fields for human motion recognition. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1969–1972. IEEE Press, New York (2008)

Author Index

J. J. Park et al. (Eds.): CUTE 2018/CSA 2018, LNEE 536, pp. 633–636, 2020.
https://doi.org/10.1007/978-981-13-9341-9