Lecture Notes in Electrical Engineering 716

James J. Park
Vincenzo Loia
Yi Pan
Yunsick Sung *Editors*

Advanced Multimedia and Ubiquitous Engineering

MUE-FutureTech 2020

Springer

Lecture Notes in Electrical Engineering

Volume 716

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, Japan, Rest of Asia

Swati Meherishi, Executive Editor (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

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

USA, Canada:

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

All other Countries:

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

**** Indexing: Indexed by Scopus. ****

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

James J. Park · Vincenzo Loia ·
Yi Pan · Yunsick Sung
Editors

Advanced Multimedia and Ubiquitous Engineering

MUE-FutureTech 2020

Springer

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

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

Vincenzo Loia
Department of Business Science
University of Salerno
Fisciano, Italy

Yunsick Sung
Department of Multimedia Engineering
Dongguk University-Seoul
Seoul, Korea (Republic of)

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-15-9308-6 ISBN 978-981-15-9309-3 (eBook)
https://doi.org/10.1007/978-981-15-9309-3

Message from the FutureTech 2020 General Chairs

FutureTech 2020 is the 15th event of the series of international scientific conference. This conference took place on August 17–19, 2020 in Jeju, Korea. The aim of FutureTech 2020 is to provide an international forum for scientific research in the technologies and application of information technology. Ever since its inception, International Conference on Future Information Technology has been successfully held as FutureTech 2019 (Xian, China), FutureTech 2018 (Salerno, Italy), FutureTech 2017 (Seoul, Korea), FutureTech 2016 (Beijing, China), FutureTech 2015 (Hanoi, Vietnam), FutureTech 2014 (Zhangjiajie, China), FutureTech 2013 (Gwangju, Korea), FutureTech 2012 (Vancouver, Canada), FutureTech 2011 (Loutraki, Greece), FutureTech 2010 (Busan, Korea, May 2010), which was the next event in the series of highly successful the International Symposium on Ubiquitous Applications & Security Services (UASS-09, USA, Jan. 2009), previously held as UASS-08 (Okinawa, Japan, Mar. 2008), UASS-07 (Kuala Lumpur, Malaysia, August, 2007), and UASS-06 (Glasgow, Scotland, UK, May, 2006).

The conference papers included in the proceedings cover the following topics: Hybrid Information Technology, High-Performance Computing, Cloud and Cluster Computing, Ubiquitous Networks and Wireless Communications, Digital Convergence, Multimedia Convergence, Intelligent and Pervasive Applications, Security and Trust Computing, IT Management and Service, Bioinformatics and Bio-Inspired Computing, Database and Data Mining, Knowledge System and Intelligent Agent, Game and Graphics, and Human-centric Computing and Social Networks. Accepted and presented papers highlight new trends and challenges of future information technologies. We hope readers will find these results useful and inspiring for their future research.

We would like to express our sincere thanks to Steering Chair: James J. (Jong Hyuk) Park (SeoulTech, Korea) and Young-Sik Jeong (Dongguk University, Korea). Our special thanks go to the Program Chairs: Giuseppe Fenza (University of Salerno, Italy), Guangchun Luo (University of Electronic Science and Technology of China, China), China Ching-Hsien Hsu (Chung Hua University, Taiwan), Houcine Hassan (Universitat Politecnica de Valencia, Spain), Jin Wang (Yangzhou University, China), Yan Li (Inha University, Korea), Ying Xia

(Chongqing University of Posts and Telecommunications, China), Qinghua Zhang (Chongqing University of Posts and Telecommunications, China), Yi Jiang (Chongqing University of Posts and Telecommunications, China), Xu Zhang (Chongqing University of Posts and Telecommunications, China), all Program Committee members, and all reviewers for their valuable efforts in the review process that 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.

FutureTech 2020 General Chairs
Kim-Kwang Raymond Choo
The University of Texas, San Antonio, USA

Jiannong Cao
The Hong Kong Polytechnic University, Hong Kong

Jungho Kang
Baewha Women's University, South Korea

Message from the FutureTech 2020 Program Chairs

Welcome to the 14th International Conference on Future Information Technology (FutureTech 2020), which will be held in Jeju, Korea on August 17–19, 2020. FutureTech 2020 will the most comprehensive conference focused on the various aspects of information technologies. It will provide an opportunity for academic and industry professionals to discuss recent progress in the area of future information technologies. In addition, the conference will publish high-quality papers which are closely related to the various theories and practical applications in multimedia and ubiquitous engineering. Furthermore, we expect that the conference and its publications will be a trigger for further related research and technology improvements in these important subjects.

For FutureTech 2020, we received many paper submissions, after a rigorous peer-review process, we accepted only articles with high quality for the FutureTech 2020 proceedings, 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!

FutureTech 2020 Program Chairs
Giuseppe Fenza, University of Salerno, Italy
Guangchun Luo, University of Electronic Science and Technology of China, China
China Ching-Hsien Hsu, Chung Hua University, Taiwan
Houcine Hassan, Universitat Politecnica de Valencia, Spain
Jin Wang, Yangzhou University, China
Yan Li, Inha University, Korea
Ying Xia, Chongqing University of Posts and Telecommunications, China

Qinghua Zhang, Chongqing University of Posts and Telecommunications, China
Yi Jiang, Chongqing University of Posts and Telecommunications, China
Xu Zhang, Chongqing University of Posts and Telecommunications, China

FutureTech 2020 Organization

Honorary Chair

Doo-soon Park, Soonchunhyang University, Korea
Guanghong Xu, Chongqing University of Posts and Telecommunications, China

Steering Chairs

James J. Park, SeoulTech, Korea
Young-Sik Jeong, Dongguk University, Korea

General Chairs

Kim-Kwang Raymond Choo, The University of Texas at San Antonio, USA
Jiannong Cao, The Hong Kong Polytechnic University, Hong Kong
Jungho Kang, Baewha Women's University, Korea

Program Chairs

Giuseppe Fenza, University of Salerno, Italy
Guangchun Luo, University of Electronic Science and Technology of China, China
China Ching-Hsien Hsu, Chung Hua University, Taiwan
Houcine Hassan, Universitat Politecnica de Valencia, Spain
Jin Wang, Yangzhou University, China
Yan Li, Inha University, Korea
Ying Xia, Chongqing University of Posts and Telecommunications, China
Qinghua Zhang, Chongqing University of Posts and Telecommunications, China
Yi Jiang, Chongqing University of Posts and Telecommunications, China
Xu Zhang, Chongqing University of Posts and Telecommunications, China

Publicity Chairs

Chao Tan, Tianjin University, China
Liang Yang, Guandong University of Technology, China
Padmanabh Thakur, Graphic Era University, India
Ping-Feng Pai, Nation Chi Nan University, Taiwan

Seokhoon Kim, Soonchunhyang University, Korea
Ling Tian, University of Electronic Science and Technology of China, China
Emily Su, Taipei Medical University, Taiwan

Workshop Chair

Jun-Ho Huh, Catholic University of Pusan, Korea

Program Committee

Homenda Wadysaw, Politechniki Warszawskiej, Poland
Zou Cliff, University of Central Florida, United States
Gonzalez Jose Antonio, University of Malaga, Spain
Chi-Fu Huang, National Chung-Cheng University, Taiwan
Buyya Rajkumar, The University of Melbourne, Australia
Qiang He, Swinburne University of Technology, Australia
Maumita Bhattacharya, Charles Sturt University, Australia
Soon M. Chung, Wright State University, USA
Alsultanny Yas, Arabian Gulf University, Bahrain
Carson Leung, University of Manitoba, Canada
Huang Kuo-Chan, National Taichung University of Education, Taiwan
Pichappan Pit, Zhejiang University, China
Caldelli Roberto, University of Florence, Italy
Salem Abdelbadeeh, Ain Shams University, Egypt
Zeyar Aung, Khalifa University, UAE
Raylin Tso, National Chengchi University, Taiwan
Wookey Lee, Inha University, Republic of Korea
Neruda Roman, Czech Academy of Sciences, Czech Republic
Sudeep Tanwar, Nirma University, India
Klyuev Vitaly, University of Aizu, Japan.

Message from the MUE 2020 General Chairs

MUE 2020 is the 14th event of the series of international scientific conference. This conference takes place on August 17–19, 2020 in Jeju, Korea. The aim of the MUE 2020 is to provide an international forum for scientific research in the technologies and application of Multimedia and Ubiquitous Engineering. Ever since its inception, International Conference on Multimedia and Ubiquitous Engineering has been successfully held as MUE-19 (Xian, China), MUE-18 (Salerno, Italy), MUE-17 (Seoul, Korea), MUE-16 (Beijing, China), MUE-15 (Hanoi, Vietnam), MUE-14 (Zhangjiajie, China), MUE-13 (Seoul, Korea), MUE-12 (Madrid, Spain), MUE-11 (Loutraki, Greece), MUE-10 (Cebu, Philippines), MUE-09 (Qingdao, China), MUE-08 (Busan, Korea), and MUE-07 (Seoul, Korea).

The conference papers included in the proceedings cover the following topics: Multimedia Modeling and Processing, Multimedia and Digital Convergence, Ubiquitous and Pervasive Computing, Ubiquitous Networks and Mobile Communications, Ubiquitous Networks and Mobile Communications, Intelligent Computing, Multimedia and Ubiquitous Computing Security, Multimedia and Ubiquitous Services, Multimedia Entertainment. Accepted and presented papers highlight new trends and challenges of Multimedia and Ubiquitous Engineering. We hope readers will find these results useful and inspiring for their future research.

We would like to express our sincere thanks to Steering Chair: James J. (Jong Hyuk) Park (Seoul Tech, Korea). Our special thanks go to the Program Chairs: Fei Hao (Shaanxi Normal University, China), Ka Lok Man (Xi'an Jiaotong-Liverpool University, China), Wei Song (North China University of Technology, China), Joon-Min Gil (Catholic University of Daegu, Korea), all Program Committee members, and all reviewers for their valuable efforts in the review process that helped us to guarantee the highest quality of the selected papers for the conference.

MUE 2020 General Chairs
Vincenzo Loia, University of Salerno, Italy
Yi Pan, Georgia State University, USA
Yunsick Sung, Dongguk University-Seoul, Korea

Message from the MUE 2020 Program Chairs

Welcome to the 14th International Conference on Multimedia and Ubiquitous Engineering (MUE 2020), which will be held at Jeju, Korea on August 17–19, 2020. MUE 2020 will the most comprehensive conference focused on the various aspects of multimedia and ubiquitous engineering. It will provide an opportunity for academic and industry professionals to discuss recent progress in the area of multimedia and ubiquitous environment. In addition, the conference will publish high-quality papers which are closely related to the various theories and practical applications in multimedia and ubiquitous engineering. Furthermore, we expect that the conference and its publications will be a trigger for further related research and technology improvements in these important subjects.

For MUE 2020, we received many paper submissions, after a rigorous peer-review process, we accepted only articles with high quality for the MUE 2020 proceedings, 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!

MUE 2020 Program Chairs
Fei Hao, Shaanxi Normal University, China
Ka Lok Man, Xi'an Jiaotong-Liverpool University, China
Wei Song, North China University of Technology, China
Joon-Min Gil, Catholic University of Daegu, Korea

MUE 2020 Organization

Honorary Chair

Young-Sik Jeong, Dongguk University, Korea
Guanghong Xu, Chongqing University of Posts and Telecommunications, China

Steering Chair

James J. Park, SeoulTech, Korea

General Chairs

Vincenzo Loia, University of Salerno, Italy
Yi Pan, Georgia State University USA
Yunsick Sung, Dongguk University, Korea

Program Chairs

Fei Hao, Shaanxi Normal University, China
Ka Lok Man, Xi'an Jiaotong-Liverpool University, China
Wei Song, North China University of Technology, China
Joon-Min Gil, Catholic University of Daegu, Korea

Publicity Chairs

Ji Su Park, Jeonju University, Korea
Jaehwa Chung, Korea National Open University, Korea
Junbo Wang, University of Aizu, Japan
Pradip Shama, SeoulTech, Korea

Workshop Chair

Jun-Ho Huh, Catholic University of Pusan, Korea

Program Committee

Rajkumar Kannan, Affiliation Bishop Heber College, India
HeonChang Yu, Korea University, Korea

Ch.Z.Patrikakis, University of West Attica, Greece
Angel D.Sappa, ESPOL Polytechnic University (Ecuador) and Computer Vision Center, Spain
Chao-Tung Yang, Tunghai University, Taiwan
Debzani Deb, Winston-Salem State University, USA
Jun-Won Ho, Seoul Women's University, Korea
Marco Cremonini, University of Milan, Italy
SungSuk Kim, Seokyeong University, Korea
Muhammad Younas, Oxford Brookes University, UK
Kwang Sik, Chung Korea National Open University, Korea
Savvas Chatzichristofis, Neapolis University Pafos, Cyprus
Pascal Lorenz, University of Haute Alsace, France
Anand Nayyar, Da Nang, Viet Nam
Sokratis Katsikas, Open University of Cyprus, Cyprus
Francisco Jose Monaco, University of São Paulo, Brazil.

Contents

Research on Optimization of Knapsack Problem in Logistics Distribution

Yan Zhang, Yanmei Zhou, and Hao Hu

Abstract In order solve the problem of low vehicle load rate effectively in modern logistics distribution, this paper proposed an adaptive optimization algorithm based on Simulated Annealing algorithm and Binary Particle Swarm Optimization algorithm. This paper tests the optimizing algorithm by building simplified models of logistics and knapsack problem model. The results of experiment show that the average distribution distance, average capacity utilization and average load rate have been improved compared to the previous results, and the number of delivery vehicles is decreased. The results indicate that the algorithm reach the purpose of optimization to knapsack problem in logistics distribution.

Keywords Logistics distribution · Load rate · Binary particle swarm · Knapsack problem

1 Introduction

At present, the algorithm of solving knapsack problem mainly include genetic algorithm [1], the particle swarm algorithm [2], the greedy algorithm, ant colony algorithm [3] and the simulated annealing algorithm et heuristic algorithms [4]. Haddar Boukthir combined the local search operators with the quantum particle swarm optimization algorithm to solve the binary knapsack problem. Glover [5] proposed a kind of improved greedy algorithm to solve the single constraint knapsack problem. Sachdeva and Goel [6, 7] proposed a kind of improved genetic algorithm to solve 0/1 knapsack problem. The algorithms mentioned above have the difficulties to take both the local search ability and global search ability into account, the algorithm will often fall into the local optimum and cannot reach the global optimal solution.

In view of the problem that the above algorithms are easy to fall into local convergence, this paper puts forward a kind of algorithm which based on simulated annealing algorithm and binary particle swarm optimization algorithm. It takes

Y. Zhang (✉) · Y. Zhou · H. Hu
Chongqing University of Posts and Telecommunications, Chongqing 400065, China
e-mail: 16458321@qq.com

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_1

the goods in express logistics and distribution as the research object as the research object. The algorithm balances the global search ability and local search ability, and then to optimize the knapsack problem in logistics distribution, in order to improve the transport vehicles load ratio and further reduce the cost of logistics and transport.

1.1 Traditional Knapsack Problem Description

Described as follows: given a knapsack with load limit M,n items, the weight of the goods $W = [w_1, w_2, \ldots, w_n]$, the corresponding value $P = [p_1, p_2, \ldots, p_n]$ of the goods, where x_j represents that whether j_{th} item is placed in the knapsack, $x_j = 0$ represent j_{th} item is not placed into the knapsack. The goal is to seek a solution that make the total value of the object in the backpack reach the largest and ensure that the total quality does not exceed the limit M of the knapsack load.

Mathematical model:

$$\text{M}ax \sum_{j=1}^{n} p_j \cdot x_j \tag{1}$$

$$s.t. \sum_{j=1}^{n} w_j \cdot x_j \leq M \tag{2}$$

$$x_j = 0 \ \ or \ \ x_j = 1 \tag{3}$$

where, $j = 1, 2, \ldots, n$, the constraint condition: $\sum_{j=1}^{n} w_j \cdot x_j \leq M$, the object function: Max $\sum_{j=1}^{n} P_j \cdot x_j$, p_j, w_j, M are all positive value.

1.2 Particle Swarm Optimization Algorithm

Particle swarm optimization algorithm is a swarm intelligence optimization according to update the iterative [8]. Composed of m particle swarm searching in a D-dimension with a certain speed. Each particle recorded best position (local optimal value) after each iteration. The best position of all the particles is global best position (global optimum). The particle's position is represented as $x_i = \left(x_{i,1}, x_{i,2}, \ldots x_{i,d}\right)$. The flight velocity of a particle is represented as $v_i = \left(v_{i,1}, v_{i,2}, \ldots v_{i,d}\right)$. The current optimal of a particle is represented as $pb_i = \left(pb_{i,1}, pb_{i,2}, \ldots pb_{i,d}\right)$; The best position of all particles is represented as $gb = (gb_1, gb_2, \ldots, gb_d)$. In each iteration process, the flight speed and position of each particle are updated based on the following formula.

$$v_{i,d}(t+1) = \omega \cdot v_{i,d}(t) + c_1 \cdot r_1 \cdot (pb_{i,d} - x_{i,d}(t)) + c_2 \cdot r_2 \cdot (gb_{i,d}(t) - x_{i,d}(t)) \tag{4}$$

$$x_{i,d}(t+1) = x_{i,d}(t) + v_{i,d}(t+1) \tag{5}$$

where, t is the number of iterations, ω is inertia weight. The second term in Eq. (4) represents its own learning ability (the ability of local search), the third term represents cooperation between particles (the ability of global search).c_1 and c_2 is acceleration coefficient,r_1 and r_2 is random numbers between 0 and1. In each iteration process, the velocity and position of the particles are updated through the Eqs. (4), (5).

2 Knapsack Problem Solution in Logistics Distribution

This paper presents an integrated optimization algorithm based on simulated annealing algorithm and binary particle swarm optimization algorithm, which can solve the knapsack problem effectively in logistics distribution. The proposed algorithm fusion simulated annealing algorithm has the ability to jump out of local optimum and binary particle swarm algorithm with the ability of easily operating and searching for the global optimal.

2.1 Simulated Annealing and Binary Particle Swarm Optimization (BPSO_SA)

This paper makes improvements to the basic binary particle swarm. The biggest drawback of basic binary particle swarm optimization algorithm is prone to premature convergence. Because the particle has the trend of close to the local in the evolutionary process, and make the search space of particles confined in a small area into a local optimum, so that the global search ability of the particles is declined. Considering the Metropolis criterion of simulated annealing algorithm that can accept the bad solution and can jump out of local optimum in the search process, so in the process of particle swarm optimization, the simulated annealing operation is used to solve the drawback that falls into the local optimum of particle swarm optimization in the search process. The proposed algorithm can balance the global and local search ability according to update the weight factor and learning factor in each iteration process. The comprehensive optimization algorithm is proposed in this paper as the specific process.

2.2 Particle Fitness Update

The search direction and speed of the particles are updating according to the fitness value of dynamic in the evolutionary process. So the particle's fitness function determines the algorithm's search performance. The model established in this paper is multi-objective function. The multi-objective function were normalized and transformed into a single objective model due to the different dimensions of each target function. Take the objectives function as the fitness value of the particle.

Vehicle loading rate:

$$\max\ Z_1 = \log\left(\frac{\sum_{j=1}^{n} w_j * x_j}{M}\right) \tag{6}$$

Vehicle volume ratio:

$$\max\ Z_2 = \log\left(\frac{\sum_{j=1}^{n} v_j * x_j}{V}\right) \tag{7}$$

Total postage:

$$\max\ Z_3 = \log\left(\sum_{j=1}^{n} val_j * x_j\right) \tag{8}$$

General objective function:

$$\max\ f(x) = Z_1 + Z_2 + Z_3 \tag{9}$$

2.3 Weight, Learning Factor Update

A larger inertia weight is beneficial to jump out of the local extreme points in search process, and a smaller inertia weight have the advantage to fast convergence and it can effectively improve the precision of search range. Learning factors c1 c2 separately adjust the local search ability and global search ability of particle. How to balance the particle's local search and global search process according to update the weight and learning factor is particularly important in the process of population evolution. This paper uses Eq. (9) to update weight and learning factor.

$$\begin{cases} \omega = \omega_{\max} - (t * (\omega_{\max} - \omega_{\min})/gen_{\max}) \\ c1 = c1_{\max} - (t * (c1_{\max} - c1_{\min})/gen_{\max}) \\ c2 = c2_{\max} - (t * (c2_{\max} - c2_{\min})/gen_{\max}) \end{cases} \tag{10}$$

where t is current evolutionary algebra, $gen_{\max}$ is total evolutionary time.

2.4 Individual Optimal Solution and the Global Optimal Solution Update

This part is about calculating current particle's fitness value fit_i. If fit_i is greater than the current particle's individual optimal solution, updating the particle's individual optimal solution. If the particle fitness values are greater than the historical global optimal value, then update the global optimal solution. If the particle fitness value is smaller than the historical global optimal value, then generate a solution randomly and calculating fitness value. With accepting the solution in the Metropolis criterion of simulated annealing algorithm, and corresponding update the global optimal solution.

2.5 Path Adjustment

Aiming at the above mentioned problems of low vehicle load rate, low capacity utilization and more vehicles numbers. This paper adjusts the path of the vehicle, and then use the BPSO_SA algorithm to make optimization for the vehicles transport loading, to maximize the total postage and reduce the unit mass's transportation cost at the same time.

Vehicle routing adjustment process: Optimizing the original path. Sort the original paths according to the total amount of the order. If total amount of orders about two adjacent paths are all less than the rated load, then integrating the two paths and using a vehicle to distribute. Detection of the whole path and adjust correspondingly until the formation of a new path that all the path's distribution volume reached vehicles rated load on this path.

3 Logistics Simulation System

This paper extracts the inherent characteristics of logistics operation and establishes a new type simulation model of fast delivery logistics system to verify the effectiveness of proposed algorithm. System mainly consists of system parameter configuration module, multi-layer urban schema generation module, order generation module,

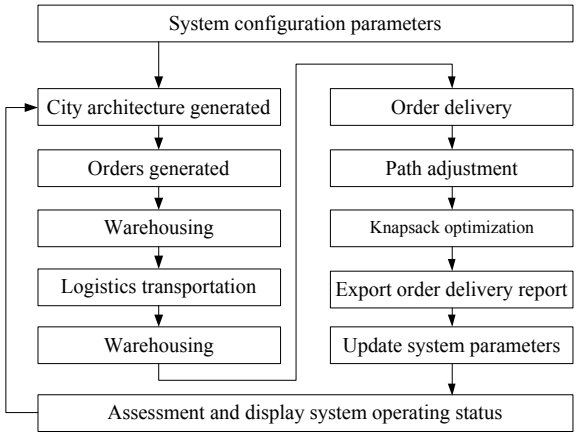


Fig. 1 Running flow chart express logistics simulation system model

order storage module, order transport module, order terminal distribution module and knapsack optimization module.et seven logistics function module.

The operating principles are defined as follows: orders are distributed from a starting point, and sending these orders to the distribution center, which is belonging to the city of order creation. Then, these orders will be delivered from distribution center to the previous level that is logistics center. After all kinds of orders have been processed and classified, sending each order to the corresponding logistics center of destination city. Next, delivering orders from logistics center to the next level that is distribution center. At last, each order will arrive to their destination through distribution center (Fig. 1).

3.1 Simulation Results and Analysis

The order of the distribution centre of assembly transport vehicle, and the BPSO-SA algorithm the loading optimization, to maximize the total postage makes the full load rate of each vehicle and loaded goods.

Randomly selecting the eleventh distribution points in January 8th to test the adjustment of the path after the optimization of the knapsack problem, test results compared to Figs. 2 and 3.

Figure 2 shows that before optimization, it uses 4 cars for delivering, the total distribution distance is 1.4493*100 km. Customer satisfaction is 2.07. After optimization, it uses 3 cars for delivering, the total distribution distance is 1.4131*100 km. Customer satisfaction is 2.20. Results are significantly improved. It is proved that the optimization algorithm is effective. When the algorithm runs in different times, the simulation data is compared with the Table 1.

The four algorithms under the same conditions were separately running for 10,20,40,60 times and the results of total mass, total postage, load ratio of the average

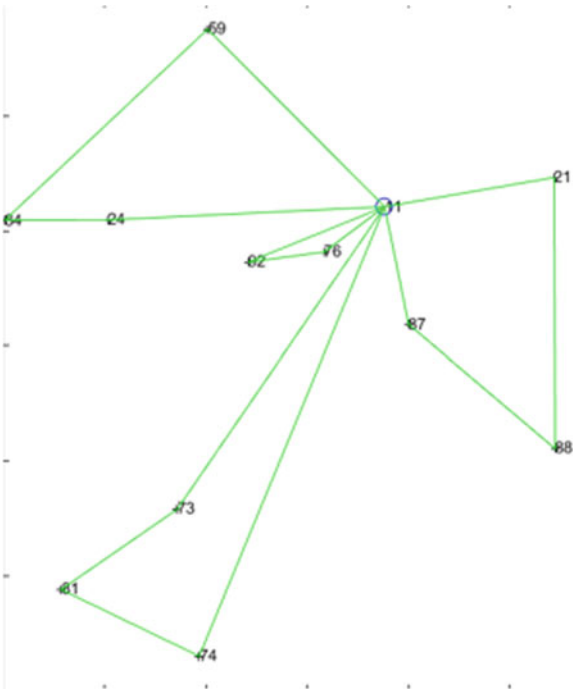


Fig. 2 Results before optimization

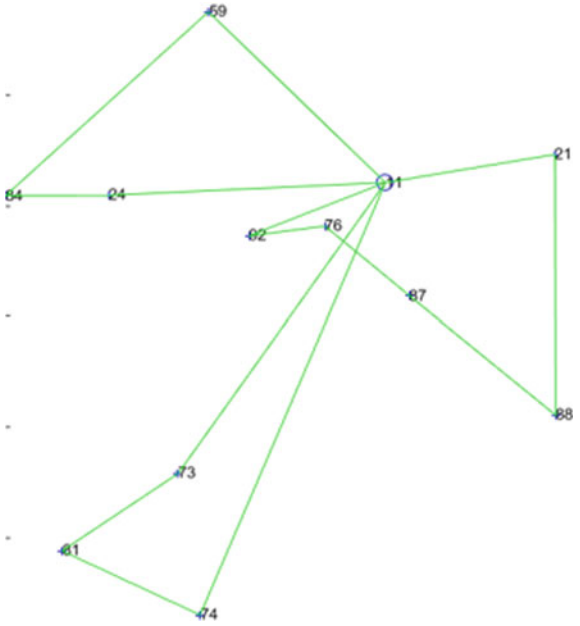


Fig. 3 Results before optimization

Table 1 Simulation results of the algorithm under different running times

次	Algorithm	Total mass			Total postage			Full-load ratio		
		Avg	Max	Min	Avg	Max	Min	Avg (%)	Max (%)	Min (%)
10	SA	597.6	600	587	5847.2	5930	5749	98.6	99.8	95.7
	BPSO	597.4	600	588	6084.4	6140	6008	98.6	99.5	97.3
	GA	593.8	600	581	5900.4	5998	5804	94.3	96.3	91.5
	BPSO_SA	598.8	600	597	6085.4	6162	6010	99.1	99.8	97.8
20	SA	596.15	600	584	5858.7	5943	5745	98.9	99.8	95.7
	BPSO	597.7	600	588	6072.9	6175	5989	98.3	99.8	96.2
	GA	594.35	600	581	5912.15	6033	5799	95.1	98.3	91.5
	BPSO_SA	598.7	600	595	6076.3	6162	5980	99.1	99.8	97.0
40	SA	596.08	600	584	5863	5993	5624	98.9	99.8	94.5
	BPSO	597.45	600	588	6068.6	6175	5977	98.5	99.8	95.3
	GA	594.13	600	570	5924.1	6070	5724	94.8	99.5	88.3
	BPSO_SA	598.78	600	594	6073.4	6162	5980	99.2	99.8	97.0
60	SA	596.1	600	575	5847.97	6011	5624	98.9	99.8	94.5
	BPSO	597.82	600	588	6068.46	6175	5977	98.5	99.8	95.3
	GA	593.64	600	570	5933.75	6128	5724	94.6	99.8	88.3
	BPSO_SA	598.84	600	594	6073	6199	5980	99.3	99.8	97.0

value, maximum value and minimum value by BPSO_SA algorithm are better than the other three algorithms (GA, BPSO, SA).

4 Conclusion

Aiming at the problem of fall into local optimum of the existing algorithms, this paper puts forward hybrid optimization algorithm based on the binary particle swarm optimization and simulated annealing algorithm. The algorithm combines better global search ability of binary particle swarm optimization algorithm and excellent local search ability of simulated annealing algorithm. And it was verified by constructing logistics simulation model and the knapsack problem model. Experimental result shows that the algorithm proposed in this paper can better optimize logistics knapsack problem and improve the transport vehicles load rate.

Acknowledgements Supported by the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJQN201900625).

References

1. Silveira JLMD, Xavier EC (2014) Pickup and delivery problem with two dimensional loading/unloading constraints. Computational Logistics. Springer, pp 31–46
2. Drake JH, Hyde M, Ibrahim K et al (2014) A genetic programming hyper-heuristic for the multidimensional knapsack problem. Kybernetes 43(9/10):1500–1511
3. Beheshti Z, Shamsuddin SM, Hasan S (2015) Memetic binary particle swarm optimization for discrete optimization problems. Inf Sci 299:58–84
4. Glover F (2013) Advanced greedy algorithms and surrogate constraint methods for linear and quadratic knapsack and covering problems. Eur J Oper Res 230(2):212–225
5. Mansour IB, Alaya I (2015) Indicator based ant colony optimization for multi-objective Knapsack problem. Proc Comput Sci 60:448–457
6. Leung SCH, Zhang D, Zhou C et al (2012) A hybrid simulated annealing metaheuristic algorithm for the two-dimensional knapsack packing problem. Comput Oper Res 39(1):64–73
7. Haddar B, Khemakhem M, Rhimi H et al (2016) A quantum particle swarm optimization for the 0–1 generalized knapsack sharing problem. Nat Comput 15(1):153–164
8. Sachdeva C, Goel S (2014) An improved approach for solving 0/1 Knapsack Problem in polynomial time using genetic algorithms. Recent Advances and Innovations in Engineering (ICRAIE). IEEE, pp 1–4

Construct of Data Compression System for Urgent Data Transmission in Smart Dust Environment

Joonsuu Park and KeeHyun Park

Abstract A smart dust environment can be useful for obtaining information on rough terrain that is difficult for humans to access. Information can be obtained by using aircraft in the Amazon rainforest to scatter fine sensors, or by using unmanned spacecraft to throw sensors on the moon's surface. Extreme environments usually have limited network performance. Given this limited network performance, it may be very inefficient to deliver or obtain all information sequentially. Therefore, we propose a method to prioritize and reduce the load by layering the measured data in the same network to solve this problem by use 2-phase methods. In the first phase, we applied the threshold delay transmission, which transmits data only when a change above the threshold occurs. This procedure is very concise and requires very little computing power, so it can be used well in a dust device. In the next phase, we apply a way of integrating the data and forwarding it to the parent node so that the matching information is shared by different data. We actually implemented two-phase methods and measured the baud rate. Where the system proposed in the experiment transmits 100% of both normal and urgent data, in the ordinary system 25.4% of normal data and 100% of urgent data are transmitted. Where the proposed system transmits 21.3% of normal data and 100% of urgent data, the ordinary system sends only 4.1% of normal data and 97.2% of urgent data. The results of these experiments validate the proposed system.

Keywords Internet of things · Smart dust · Remote transmit · Monitoring system

1 Introduction

The Internet of Things (IoT) is a close connection between the physical and digital worlds [1–4]. The IoT, which conceptually and experimentally began to be

J. Park · K. Park (✉)
Department of Computer Engineering, Keimyung University, Daegu, South Korea
e-mail: khp@kmu.ac.kr

J. Park
e-mail: parkjoonsuu@gmail.com

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_2

studied only a few years ago, has already been applied in several sectors, such as transportation, smart city, smart domotics, smart health, e-governance, assisted living, e-education, retail, logistics, agriculture, automation, industrial manufacturing, and business/process management, etc., and we are already benefiting from this technology, directly or indirectly [3, 5, 6].

Smart Dust technologies is a field of IoTs in a broad sense, involving very small dust-sized airborne sensors for monitoring physical spaces such as buildings, roads, clothes, and the human body, gathering information such as temperature, humidity, acceleration, and pressure. Smart Dust is a technology that can detect and manage a wireless network [7–11]. Smart Dust has the advantage of being able to collect and disperse large amounts of information scattered around, but it is difficult to arrange and manage when installing in inaccessible areas.

Smart Dust environments can be useful for obtaining information on rough terrain that is difficult for humans to access. Information can be obtained by using aircraft in the Amazon rainforest to scatter fine sensors, or by using unmanned spacecraft to throw sensors on the moon's surface. Extreme environments usually have limited network performance. Given this limited network performance, it may be very inefficient to deliver or obtain all information sequentially. Therefore, we propose a method to prioritize and reduce the load by layering the measured data in the same network to solve this problem.

The remainder of this paper is organized as follows: This chapter introduces previous studies and the data used in the research. Chapter "Entity Summarization in Fuzzy Knowledge Graph Based on Fuzzy Concept Analysis" introduces a hierarchical transmission method designed to achieve our goal. In Sect. 4, we validate the proposed algorithm and conclude in Sect. 5.

2 Previous Work

Ghobakhlou, A., S. Shanmuganthan, and P. Sallis proposes a framework for managing data using Wired Sensor Networks (WSN) [12]. They divide the layers of the system into three (Mote, Server, and Application) and discuss how to provide a good framework to serve information to the application layer. They also create an excellent environment for providing services by isolating the application layer from the mote layer. [13] also designed their framework nicely. They focused on smoothly servicing by separating the layers, similar to [12]. They created an interface layer between the database layer and the sensing layer. The interface layer communicates with the application data through the web, and with the database layer through a database connection. The feature of both frameworks described above is that connectivity on a scale that is imaginable is assumed. The same is true for [14]. It considers more about the hardware part of the device than [12] and [13]. However, this consideration is for one device. In other words, [14] does not consider an enormous number of environments, such as a smart dust environment.

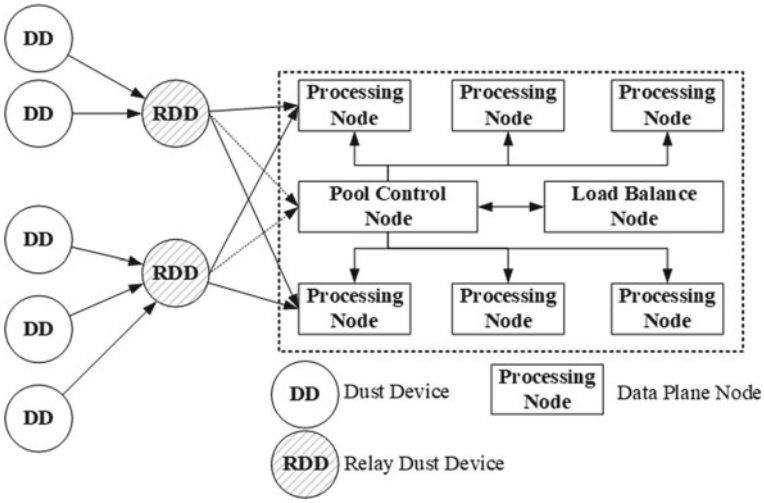


Fig. 1 An overview of physical devices in a smart dust environment

Figure 1 shows the configuration of the physical devices [15–17]. Figure 1 consists of a multi-layered form that we can easily imagine. The only difference is the difference between the Dust Device (DD) and the Relay Dust Device (RDD), which can be connected in huge numbers.

We classify physical devices into devices and nodes. A device is a consumable device that has very limited computing power on a node, and is very inexpensive. Nodes are devices with high computing power that can handle large amounts of data and can do more if required by the administrator, and there are only a limited number in the system.

DD is a sensing node called Dust Device. The role played by this node is actually to be sprayed on fields such as rough terrain, mountains, etc. to collect the surrounding climate and environmental information. If the type of DD is changed, the system is changed to perform a different role. In other words, DD is a very important node that determines the role of the system as a whole.

RDD is a node called Relay Dust Device. As can be inferred from the name and Fig. 1, this node plays the role of DD sensing around information, and collects and transmits information from other peripheral DDs. This also means that RDD requires higher computing power than DD. The relay requires better network performance, and also requires caching. The two kinds of physical devices mentioned above are "devices."

3 Hierarchical Transmit Method

The smart dust environment is one with vast quantities of sensors scattered everywhere. Depending on the purpose of the system, there are many different types of sensors. With huge numbers of various types of sensors scattered in an environment and large amounts of information gathered, inevitably a bottleneck results, in which data is concentrated at intermediate and back end nodes. The range of delay tolerance of the data depends on the type of sensor that collected it. For example, data whose sensor type is not sensitive to delays, such as thermometers and hygrometers, will have little effect on the system or human life, even if delayed. But data that can cause major problems when not handled urgently, such as information regarding an earthquake, is very sensitive to delay and must be sent immediately.

As described above, when a large amount of data causes a bottleneck, even data that requires urgent transmission cannot utilize network resources, and as a result, "urgent" transmissions are missed, causing various problems.

We assume a system in a normal environment where some types of data allow delays and some other types do not allow delays. In a typical system, a unique feature of sensing data called "delay tolerance" is used to implement a system that smoothly transmits data requiring urgent transmission and minimizes the delay of data that allows delay. To use this, we classify data into normal data which can be delayed slightly and urgent data which does not allow delay. Urgent data cannot be integrated, because it must be sent as soon as it is sensed.

Therefore, we propose a method of eliminating bottlenecks by freeing up network resources available to urgent data through reducing and integrating them without affecting the normal data size system. This method consists of two-phase (i.e. the 1st-phase and the 2nd-phase).

3.1 The Normal Data Threshold Filtering Phase (The First Phase)

Urgent data is mostly accompanied by the possibility of major damage. Normal data can sometimes be accompanied by similar damage as urgent data.

Normal data with significant damage is accompanied by large data changes. For example, temperature data are not significant in most cases, but if changes in temperature are due to fire, etc. this can be accompanied by damage. Conversely, normal data that do not involve abrupt temperature changes are less likely to involve damage. Therefore, we applied the threshold delay transmission, which transmits data only when a change above the threshold occurs. This procedure is very concise and requires very little computing power, so it can be used well in DD.

Table 1 shows temperature, humidity, sunshine, and ground temperature data collected on May 1, 2019, in Daegu, South Korea.

Table 1 Sensing data

Time	Temperature	Humidity	Temperature ($\Delta = 1$)	Humidity ($\Delta = 1$)
2019–05-01 0:00	12.6	94	12.6	94
2019–05-01 1:00	12	96		96
2019–05-01 2:00	12.1	97		97
2019–05-01 3:00	11.6	97	11.6	
2019–05-01 4:00	10.8	98		98
2019–05-01 5:00	10.7	98		
2019–05-01 6:00	10.9	98		
2019–05-01 7:00	12.1	98		
2019–05-01 8:00	14.7	84	14.7	84
2019–05-01 9:00	17.3	57	17.3	57
2019–05-01 10:00	18.4	49	18.4	49
2019–05-01 11:00	18.9	45		45
2019–05-01 12:00	21.1	43	21.1	43
2019–05-01 13:00	21.9	37		37
2019–05-01 14:00	22.7	34	22.7	34
2019–05-01 15:00	23.4	30		30
2019–05-01 16:00	23.4	29		29
2019–05-01 17:00	23.2	29		
2019–05-01 18:00	22.2	30		30
2019–05-01 19:00	21.5	26		26
2019–05-01 20:00	20.3	26	20.3	
2019–05-01 21:00	19.1	30		30
2019–05-01 22:00	18.7	35	18.7	35
2019–05-01 23:00	17.8	44		44

In Table 1, the data in the "Temperature" column of the row with the Time column "2019–05-01 0:00" is 12.6, and the data in the "Temperature" column of the row with the Time column "2019–05-01 1:00" is 12. The two data have a difference of 0.6. If the threshold is 1, the data "2019–05-01 1:00" is not sent. Therefore, setting the threshold on which integration is based is very important. The threshold must be set differently for temperature, but 1 is a very large change in the sunshine.

The two columns with Δ in Table 1 show the unified two columns of data. Δ is a variable that specifies the range of data to be integrated. Δ should be chosen carefully because it is closely related to the ratio of data compression, and at the same time is also closely related to data reliability. The larger the Δ, the higher the compression ratio, but the lower the reliability; the lower the Δ, the higher the reliability, but the lower the compression ratio. That is, Δ must be determined carefully because it determines the trade-off between reliability and compression ratio.

The integrated sensing data has the effect of reducing 88 pieces of data to 46 (approximately 48% reduction), although the Δ is roughly set and has the limitation of using data from a specific date.

3.2 The Normal Data Integration Phase (The Second Phase)

The normal data threshold filtering phase was a procedure to reduce the number of pieces of data transmitted in DDs. The normal data integration phase is a process of integrating the data when the data arrived at the RDD to the Group of Processing Nodes. In a hierarchical network, data from child nodes inevitably gathers into parent nodes. This means that data from several DDs are gathered into one RDD. When the data gathered is sent to the higher layer nodes in the RDD, much of the same information is included. For example, sensing data, sensing device ID, sensing location, etc. may be different, but information such as type of sensing information and RDD passing through may match. Therefore, we may apply a way of integrating the data and forwarding it to the parent node so that the matching information is shared by different data.

4 Experiment

We collected real data from the Korea Meteorological Administration [18] and the Korea National Information Society Agency [19] to verify that the proposed 2-phase algorithm actually eliminates bottlenecks and increases system throughput.

As shown in Table 2, data of 1,218 bytes is reduced by about 46% (658 bytes) when 1-phase is executed. If we run 2-phase alone, you can see that the data size is reduced by about 35% (792 bytes). As a result, the result of passing both 1 and 2-phase reduced the size by about 61% (472 bytes).

We compared the throughput with the ordinary system to verify that the network resources are effectively used proportionally to the reduced size. Table 2 shows the amount and rate of data processed for 10 and 60 s. In the 60-s throughput test, both processed 100% urgent data. However, the proposed system processed 100% of normal data, whereas the ordinary system processed only 25.4% of normal data. This is a huge difference considering that the normal data accounted for 90% of

Table 2 The size of the data in phases

	Ordinary	1-phase	2-phase	1-phase 2-phase
Size (bytes)	1218	658	792	472
The ratio of ordinary (%)	100	54	65	39

the total data. Throughput measurement experiments performed for 10 s show a more pronounced difference. The proposed system sent 100% of urgent data and processed 21.3% of normal data, whereas the ordinary system not only processed 4.1% of normal data but also did not send all urgent data (sent 97%).

5 Conclusion

We proposed a two-phase method to resolve the bottleneck that adversely affects data transmission in a typical system where both urgent and non-urgent (normal) data are mixed.

The normal data threshold filtering phase reduces data size and the number of transmissions by filtering data when the information collected from the DD does not exceed the threshold. In the normal data integration phase, the information transmitted from the RDD eliminates duplication by reducing the data size.

In addition, where the system proposed in the experiment transmits 100% of both normal and urgent data, in the ordinary system 25.4% of normal data and 100% of urgent data are transmitted. Where the proposed system transmits 21.3% of normal data and 100% of urgent data, the ordinary system sends only 4.1% of normal data and 97.2% of urgent data. The results of these experiments validate the proposed system.

Acknowledgements 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. Atzori GL, Iera A, Morabito G (2010) The internet of things: a survey. Comput Netw 54(15):2787–2805
2. Gonzalez GR, Organero MM, Kloos CD (2008) Early infrastructure of an internet of things in spaces for learning. In: 2008 eighth IEEE international conference on advanced learning technologies, pp 381–383
3. Ray PP (2018) A survey on Internet of Things architectures. J King Saud Univ Comput Inf Sci 30(3):291–319
4. Sterling B (2005) Shaping things–mediawork pamphlets. The MIT Press, Cambridge
5. Gubbi J, Buyya R, Marusic S, Palaniswami M (2013) Internet of Things (IoT): a vision, architectural elements, and future directions. Future Gener Comput Syst 29(7):1645–1660
6. Miorandi D, Sicari S, De Pellegrini F, Chlamtac I (2012) Internet of things: Vision, applications and research challenges. Ad Hoc Netw 10(7):1497–1516
7. Warneke B, Last M, Liebowitz B, Pister KS (2001) Smart dust: communicating with a cubic-millimeter computer. Computer 34(1):44–51
8. Kahn JM, Katz RH, Pister KS (2000) Emerging challenges: mobile networking for "smart dust". J Commun Netw 2(3):188–196

9. Kahn JM, Katz RH, Pister KS (1999) Next century challenges: mobile networking for "Smart Dust". In: 5th annual ACM/IEEE international conference on mobile computing and networking, 271–278
10. N. I. I. P. Agency (2004) Smart dust. [Online] Available: https://www.itfind.or.kr/WZIN/jugidong/1140/114004.htm. Accessed 01 July 2018 (in Korean)
11. NICCOLAI, Lorenzo et al (2019) A review of smart dust architecture, dynamics, and mission applications. Progr Aerosp Sci
12. Ghobakhlou A, Shanmuganthan S, Sallis P (2009) Wireless sensor networks for climate data management systems. In: 18th IMACS world congress-MODSIM09 international congress on modelling and simulation
13. Kanagaraj E et al (2015) Cloud-based remote environmental monitoring system with distributed WSN weather stations. In: 2015 IEEE sensors
14. Susmitha P, Sowya Bala G (2014) Design and implementation of weather monitoring and controlling system. Int J Comput Appl 97(3)
15. Pak J, Park K (2018) A high-performance implementation of an IoT system using DPDK. Appl Sci 8(4):550
16. Park J, Park K (2019) A study on system architecture design for plane dynamic scaling in smart dust environments. In: 2019 The KIPS Spring conference, 0718–0720 (in Korean)
17. Park K, Kim I, Park J (2018) A high speed data transmission method for DPDK-based IoT systems. In: International conference on future infomation & communication engineering, 325–327 (in Korean)
18. KMA (2018) The Korea meteorological administration [online], https://data.kma.go.kr/. Accessed 01 Sep 2018
19. NISA (2018) National information society agency [online], https://www.data.go.kr. Accessed 01 Sep 2018

Entity Summarization in Fuzzy Knowledge Graph Based on Fuzzy Concept Analysis

Erhe Yang, Fei Hao, Jie Gao, and Doo-Soon Park

Abstract Considerable attention has recently been devoted to Knowledge Graph (KG), which has been applied in many domains. However, the information is often imprecise and vague when constructing the knowledge graph and thus the Fuzzy Knowledge Graph (FKG) emerged. Considering the increasing data in FKG, this paper firstly formulates the entity summarization in FKG and proposes an approach leveraging Fuzzy Formal Concept Analysis (FFCA). More specifically, the predicates and objects in RDF triples are deemed as attributes and objects in FFCA, respectively. Then, the fuzzy formal context can be obtained and the fuzzy concept lattice can be constructed. Finally, the concepts are ranked by the cardinality of the extent in concept lattice and the vague value of objects in RDF triples.

Keywords Entity Summarization · Fuzzy Knowledge Graph · Fuzzy Formal Concept Analysis

This research was supported by the National Natural Science Foundation of China (Grant No. 61702317), and the National Research Foundation of Korea (NRF-2020R1A2B5B01002134) and was also supported by the Natural Science Basic Research Plan in Shaanxi Province of China (2019JM-379). This is also a part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 840922. This work reflects only the authors' view and the EU Commission is not responsible for any use that may be made of the information it contains.

E. Yang · F. Hao (✉) · J. Gao
School of Computer Science, Shaanxi Normal University, Xi'an, China
e-mail: feehao@gmail.com

F. Hao
Department of Computer Science, University of Exeter, Exeter, UK

D.-S. Park (✉)
Department of Computer Software Engineering, Soonchunhyang University, Asan, South Korea
e-mail: parkds@sch.ac.kr

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_3

1 Introduction

The term Knowledge Graph (KG) was coined by Google in 2012 and KG has been successfully applied in many fields, such as entity summarization [1], entity alignment [2], etc. KG describes entities and their relationships by employing the Resource Description Framework (RDF)-style triples. RDF datasets are collections of subject-predicate-object triples to describe the resources and their relationships on the web. However, the relationships between entities and their properties are often imprecise and uncertain. Besides, the descriptions of objects are accompanied by a degree of uncertainty. Thus, it is necessary to build the Fuzzy Knowledge Graph (FKG). The literature [3] attempts to construct a fuzzy knowledge base from a set of documents. Xiong et al. [4] proposed a general framework to construct fuzzy knowledge bases using feature selection. This framework can reduce the computational costs of fuzzy knowledge base construction. Nevertheless, users may be overwhelmed by the massive entity descriptions. It is hard for the end-users to search for the most important information from the lengthy FKG. As the information comes from various sources, it is often vague or ambiguous. The constructed FKG is labeled with vague value, expressing that the confidence level of the information.

Example 1 Figure 1 shows a toy example of the entity in a fuzzy knowledge graph. The entity is selected from the datasets in [5]. We assume that the objects of entity are ambiguous, thus we can add the possible values as shown in the brackets. The value within the scope of [0, 1] in the bracket represents the confidence level of the objects. People are often confused with the FKG, due to the excessive volume of descriptions of the entity. Thereby, it is essential to tackle this issue.

Nevertheless, existing works only focus on the entity summarization in KG rather than in FKG. Therefore, this paper firstly formalizes the problem of entity summarization in FKG. Formal concept analysis (FCA) is a useful mathematical theory that modeling the generalization and generalization relation between objects and attributes. As its generalization, Fuzzy Formal Concept Analysis (FFCA) [6] aims

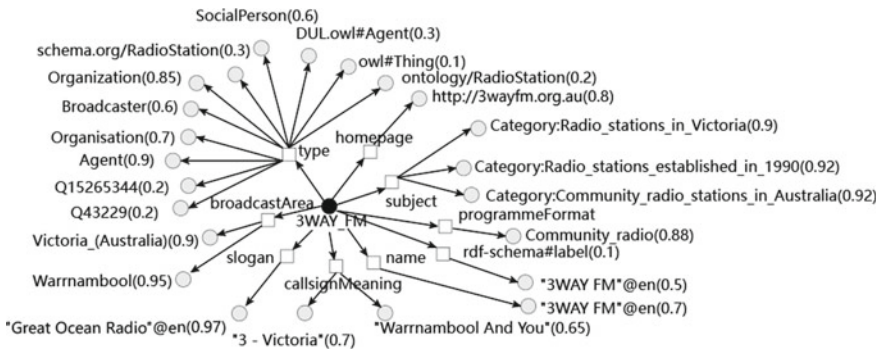


Fig. 1 The entity "3WAY_FM" in fuzzy knowledge graph

at processing fuzzy sets and representing conceptual knowledge and hierarchies. To address the novel problem of entity summarization in FKG, we present an entity summarization approach based on FFCA. Specifically, the predicates and objects of the entities in FKG are converted into attributes and objects in fuzzy formal context. Then, the fuzzy formal concepts are obtained and ranked by the weight of concepts and the vague degree of objects in RDF triples. The weight of concepts is based on the cardinality of objects in extent of concepts, which is inspired by [5]. Finally, the ranked RDF triples whose objects labeled with vague value can be obtained.

The contribution of this paper is twofold:

- Problem Formalization of Entity Summarization in FKG: We pioneer the formalization of entity summarization in FKG, which aims to select the most important and compact information of the entity in FKG. This paper focuses on the situation where the object of triples labels with vague value. The goal of this paper is to obtain the top-k RDF triples from the lengthy entity descriptions.
- FFCA-based Entity Summarization Approach in FKG: The central idea of the proposed approach is that the less the cardinality of objects in a fuzzy concept is, the more important the fuzzy concept is. Concretely, given a fuzzy knowledge graph R, we construct the fuzzy formal context K in which the tokenized objects and the predicates in fuzzy RDF triples are considered as objects and attributes while the vague degree of objects serves as the value of K. Then, the fuzzy formal concept lattice L can be built and the fuzzy concepts can be ranked by the cardinality of extent in concepts and the vague degree of objects in RDF triples. Finally, the top-k triples can be obtained.

2 Related Work

For the research on entity summarization in KG, RELIN [7] employs a variant of the random surfer model to obtain the summarized triples. Gunaratna et al. [1] presented an entity summarization approach FACES that incorporating diversity, uniqueness, and popularity of the facts. Their approach utilizes the clustering algorithm Cobweb to select the representative facts. In the literature [8], the problem of context-aware entity summarization is formalized and a Personalized Page Rank algorithm is adopted by considering user preference. With respect to the summarization approach using Formal Concept Analysis, Kim et al. [5] proposed an entity summarization approach KAFCA based on FCA. KAFCA regards the objects and predicates of entities as objects and attributes in formal context, respectively. Then it ranked the RDF triples according to the weights of extents of concepts in concept lattice. The work [9] presents a query-focused twitter summarization framework for obtaining the most representative tweets for users and summarizes them based on FFCA. However, there is no existing study focusing on entity summarization for the fuzzy knowledge graph. Hence, we present a solution for this novel problem.

3 Proposed Approach

This section firstly provides the definitions about Fuzzy Knowledge Graph and Fuzzy Concept Analysis in Sect. 3.1. Then the problem formulation is given in Sect. 3.2. Finally, Sect. 3.3 details the proposed approach.

3.1 Fuzzy Knowledge Graph

Definition 1 (Fuzzy RDF data graph) [9]. A fuzzy RDF data graph G is represented by a 6-tuple (V, E, Σ, L, μ, ρ), where V is a finite set of vertices, $E \subset V \times V$ is a set of directed edges, Σ is a set of labels, $L: V \cup E \rightarrow \Sigma$ is a function assigning labels to vertices and edges, $\mu: V \rightarrow [0, 1]$ is a fuzzy subset, and $\rho: E \rightarrow [0, 1]$ is a fuzzy relation on the fuzzy subset μ. Note that $\rho(v_i, v_j) \leq \mu(v_i) \wedge \mu(v_j)$, where $v_i, v_j \in V$.

Definition 2 (Fuzzy Formal Context) [10]. A Fuzzy Formal Context is a triple $K = (O, A, R = \varphi(O \times A))$, where O is a set of objects, A is a set of attributes, and R is a fuzzy relation in $O \times A$. Each pair $(o, a) \in R$ has a membership value $\mu(o, a)$ in [0,1].

Definition 3 (Fuzzy Formal Concept) [10]. Given a fuzzy context $K = (O, A, R = \varphi(O \times A))$, a confidence threshold T, and two sets E, I, such that $E \subseteq O$ and $I \subseteq A$, consider the dual sets E‘ and I‘, defined, respectively as follows:

$E' = \{a \in A | \mu(o, a) \geq T \; \forall o \in E\}$.

$I' = \{o \in O | \mu(o, a) \geq T \; \forall a \in I\}$.

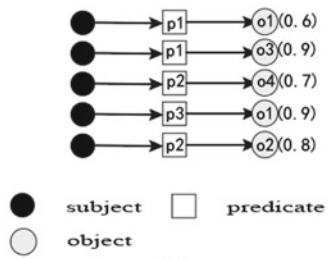
The sets E and I are referred to as the extent and the intent of the fuzzy concept, respectively.

3.2 Problem Description

Considering the increasing scale of data in FKG, it is necessary to provide a summary of the numerous descriptions of the entity. Entities in FKG are represented by RDF triples labeled with vague values. Entity summarization in FKG aims at providing a compact summary from lengthy entity descriptions. Given a fuzzy knowledge graph R, the entity summarization targets at selecting top-k RDF triples.

3.3 Proposed Approach

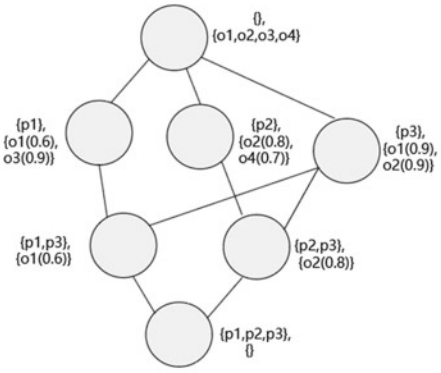
Inspired by the approach in [5], we proposed a solution to tackle the problem of entity summarization in FFG. More specifically, as shown in Fig. 2a is the RDF



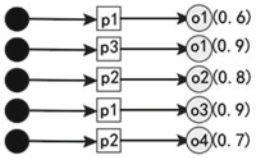
(a)

	predicate 1	predicate 2	predicate 3
tokenized object 1	0.6		0.9
tokenized object 2		0.8	0.9
tokenized object 3	0.9		
tokenized object 4		0.7	

(b)



(c)



(d)

Fig. 2 The framework of our proposed approach

triple whose objects labeled with vague values. Figure 2b is the fuzzy formal context and the corresponding fuzzy concept lattice is shown in Fig. 2c. The output RDF triples are shown in Fig. 2d. To be more precise, the ranking principle is that we firstly rank the triples according the cardinality of the extent in concepts and then by the vague values of objects when the cardinalities of objects are same. In practice, the results can be pruned according to the vague value. For example, we can remove the triples with vague value less than 0.7. Therefore, some less believable information can be trimmed (Fig. 3).

Adopting the proposed approach for the entity in Fig. 1, the following summary results can be obtained.

4 Conclusion

This paper presents an approach of entity summarization in FKG employing FFCA. Concretely, we firstly convert the predicates and objects of entity in RDF triples into attributes and objects in FFCA, respectively. Then the fuzzy formal context

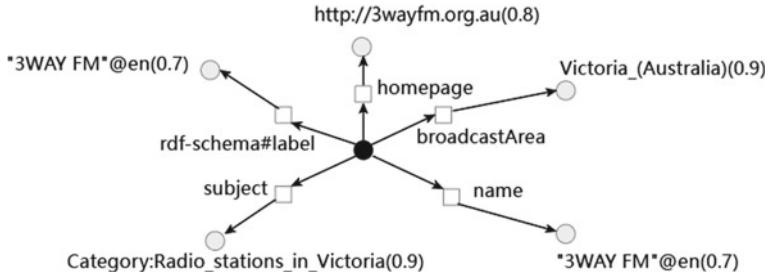


Fig. 3 The corresponding summarization of the entity "3WAY_FM" in fuzzy knowledge graph

can be obtained and the fuzzy concepts can be built. Finally, we rank the fuzzy concepts by the hierarchy of concepts in fuzzy concept lattice and the vague value of objects in RDF triples. In the future, comprehensive experiments will be conducted for evaluating the performance of the proposed approach.

References

1. Gunaratna K, Thirunarayan K, Sheth A (2015) FACES: diversity-aware entity summarization using incremental hierarchical conceptual clustering. In: Twenty-Ninth AAAI conference on artificial intelligence
2. He F, Li Z, Qiang Y, Liu A, Liu G, Zhao P, Chen Z (2019) Unsupervised entity alignment using attribute triples and relation triples. In: International conference on database systems for advanced applications. Springer, Cham, pp 367–382
3. Yarushkina N, Moshkin V, Filippov A, Guskov G (2018) Developing a fuzzy knowledge base and filling it with knowledge extracted from various documents. In: International conference on artificial intelligence and soft computing. Springer, Cham, pp 799–810
4. Xiong N, Funk P (2005) Construction of fuzzy knowledge bases incorporating feature selection. Soft Comput 10(9):796–804
5. Kim EK, Choi KS (2018) Entity summarization based on formal concept analysis. In: 1st international workshop on entity retrieval, EYRE
6. Tho QT, Hui SC, Fong ACM, Cao TH (2006) Automatic fuzzy ontology generation for semantic web. IEEE Trans Knowl Data Eng 18(6):842–856
7. Cheng G, Tran T, Qu Y (2011) Relin: relatedness and informativeness-based centrality for entity summarization. In: International semantic web conference. Springer, Berlin, Heidelberg, pp 114–129
8. De Maio C, Fenza G, Loia V, Parente M (2015) Online query-focused twitter summarizer through fuzzy lattice. In: 2015 IEEE International conference on fuzzy systems (FUZZ-IEEE). IEEE, pp 1–8
9. Li G, Yan L, Ma Z (2019) An approach for approximate subgraph matching in fuzzy RDF graph. Fuzzy Sets Syst 376:106–126
10. Formica A (2012) Semantic web search based on rough sets and fuzzy formal concept analysis. Knowl Based Syst 26:40–47

Interlacing Data to Classify Software in Linear Regression Approach

Hyun-il Lim

Abstract Linear regression is widely used in prediction problems by modeling the relationship between input and output data. In this paper, we propose an approach to interlacing data to classify similar software in the linear regression approach. Multiple linear regression models are generated from the training of the interlaced feature data. By combining the multiple models, we can apply the proposed approach in classifying similar software. Experiments are performed to evaluate the proposed approach as compared to the conventional linear regression. The experimental results show that the proposed method can classify similar software more accurately compared to the conventional linear regression model. The proposed approach is expected to be an effective approach to improve the accuracy of linear regression and it is expected to be applied in various areas of prediction problems.

Keywords Machine learning · Linear regression · Software classification · Data analysis

1 Introduction

With the rapid development of computing power in recent information systems, many real-life problems are being solved by using machine learning approaches [1–3]. Unlike traditional algorithm-based problem-solving methods, machine learning methods generate models that can solve the specific problems on their own through learning from training data. To generate a model for solving problems on its own, it is important to effectively utilize training data. In this paper, through an approach to interlacing training data, we propose a method of applying linear regression in classifying software effectively.

Linear regression [1, 4] is a method of machine learning that can predict results by modeling the relationship between input data and output results. Generally, multi-dimensional data is used as input data and used to find a model for representing the

H. Lim (✉)
Department of Computer Engineering, Kyungnam University,
Changwon, Gyeongsangnam-do 51767, South Korea
e-mail: hilim@kyungnam.ac.kr

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_4

relationship with the output of the input data. In this paper, a method is proposed to interlace the data from the features of input data to improve the performance of classifying similar software in the linear regression approach. The interlaced features of data are used to train and generate models to analyze and classify software. The proposed approach is evaluated with a set of benchmark data from Java applications.

2 Linear Regression in Machine Learning

In this section, the concept of linear regression in machine learning, and the basic approach of this paper is presented.

2.1 Linear Regression in Software Classification

Linear regression is one method among statistical approaches to analyze data, and it is effectively utilized as a method of machine learning. The linear regression approach can analyze training data and generate a model for describing the relationship between input data and output. Therefore, linear regression is a method of finding linear relationships between independent and dependent variables through data analysis. This method can model relationships between multiple independent variables and output values. For independent variables $[x_1, \ldots, x_k]$ and the dependent variable y the approach decide a relationship between $[x_1, \ldots, x_k]$ and y from the features of training data. The generated model is used to predict output y from input features of data $[x_1, \ldots, x_k]$. So, the linear model generated from the linear regression approach is described as follows:

$$\gamma = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon,$$

where the β_i are regression coefficients for representing the relationship between ith input data x_i and output y and the combination of the coefficients can construct the linear model between input data x_i and the output result y. So, the regression coefficients β_i are adjusted through the process of learning from training data. After constructing the prediction model for input and output, we can apply the problem to predict output results from input data.

Figure 1 shows an example of applying linear regression to one-dimensional input data x and output y. To construct an effective model for output from inputs, the model should minimize errors on all the training data to reduce the errors in predicting results from new input data. To minimize the errors of the results, we can generally use mean squared error to the training data to construct the model so that the predictive model can minimize the sum of errors for all training data. Thus, in Fig. 1, the line that

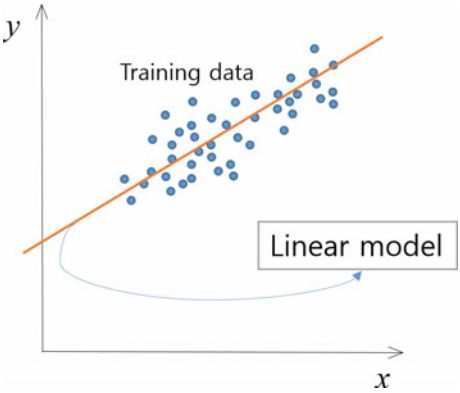


Fig. 1 An example of linear regression for input data x and output result y

minimizes errors for all data can be used to predict the relationship between inputs and outputs with the minimum errors.

2.2 The Basic Approach of Interlacing Data

In linear regression, input data consist of several features of the input. Therefore, the data are expressed in multidimensional data, and each dimension represents the unit information that describes the features of the data. In the linear regression model, these multiple feature data all affect the results that are predicted by the model. However, some feature values among all the features may contain noise values that hinder estimating the correct results in the linear regression model. So, if we can properly remove these noise feature values in training linear regression models, we can improve the accuracy of the model in predicting results with the model.

However, it may take a lot of time and effort to further analyze feature data to eliminate noise accurately before applying linear regression. Because each feature data has unique information of input data, it is difficult to differentiate and remove noise values that may reduce the accuracy of prediction. To mitigate this difficulty, we propose a method to interlace input data by extracting partial data of the total feature data. The conventional linear regression model uses all of the input data for training models, but the proposed method only uses some of the data alternately extracted from the input data.

Figure 2 illustrates an example of interlacing input data proposed in this paper. The original input data consists of several feature values. To interlace the input data, we alternatively extracted one-third of the features to construct an interlaced input data. By extracting feature values multiple times, we can construct a set of interlaced data to reduce the effect of noise features in training a linear regression model.

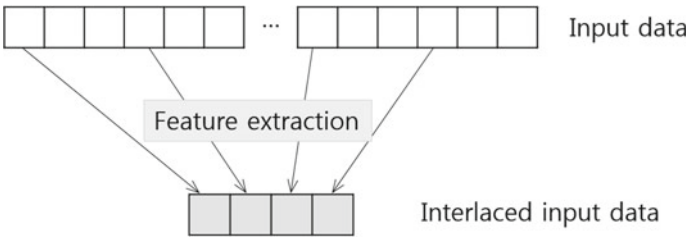


Fig. 2 An example of extracting features from original data to generate interlaced input data

3 Interlacing Data in Linear Regression

When feature values are extracted from the training data of linear regression, only a fraction among the total feature data is applied to training and generating a linear regression model. Therefore, if only a single model is trained with the feature extracted from the training data, the feature values excluded from the training will not be reflected in the results that are predicted by the model. To mitigate these shortcomings, we propose an approach to apply a set of training data that are interlaced from the original input data alternatively. So, interlacing features of data can generate multiple linear regression models. By interlacing data in linear regression, even if some models contain feature data that have a noise for predicting correct results, the noise features can be effectively suppressed by combining multiple models. Therefore, the combination of multiple linear regression models from interlaced features of input data can predict more reliable results as compared to a single model using the overall data.

Figure 3 shows an approach to interlacing data by extracting features from the original input data in linear regression. The whole input data is divided into several

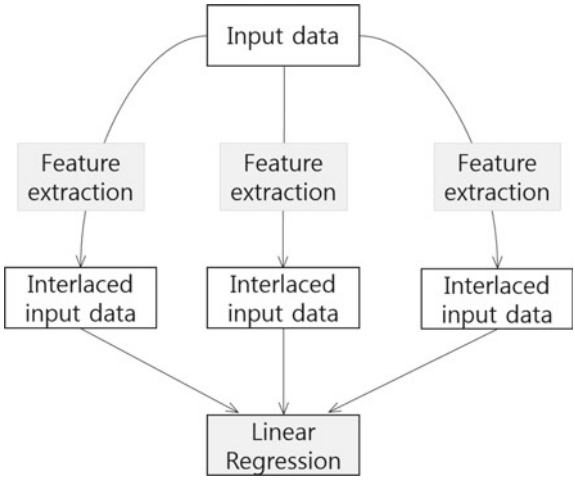


Fig. 3 An approach to interlacing data by extracting features in linear regression

feature data by alternatively interlacing the original data. The interlaced feature data are used as input data to train and generate linear regression models for the data. Each generated model can be used as a predictor alone. The overall result can be predicted through a combination of results that are generated from the predictors of the interlaced input data. Because multiple models are generated from the interlaced input data that are extracted from the original data, a combination of the models can produce more reliable results as compared to a single predictive model generated by the conventional linear regression.

4 Experimental Evaluation

In this paper, we have proposed an approach to interlacing feature from input data in the linear regression to classify similar software. To evaluate the proposed approach, we have implemented the approach with Python [5] and scikit-learn library [6] for classifying similar software. The experiment is performed in Microsoft Windows 7 with 16 GB main memory. We used a benchmarking software set for training and evaluation, independently. A Java application, ANTLR [7], is used as a training data set for building linear regression models for interlaced features from the training data. To reflect similar versions of software for classification, we have used the Smokescreen obfuscator to modify and generate similar versions of the original benchmark software. For evaluation of the approach, we have generated a test data set from Jakarta ORO [8] to evaluate the proposed approach. Table 1 shows the benchmarking set for the experiment. The numbers of data for training and testing are 13,689 and 2,252, respectively.

Table 1 The benchmarking set for training and testing linear regression models for interlaced data presented in Sect. 3

	Training data	Test data
Benchmark software name	ANTLR	Jakarta ORO
Number of data	13,689	2,500

Table 2 The experimental results of linear regression for interlaced data as compared to the general linear regression model

	Conventional linear regression	Linear regression for interlaced data
Number of interlaced data	–	3
Number of total classification	2,500	2,500
Number of correct classifications	2,252	2,356
Overall accuracy	90.08%	94.24%

Table 2 shows the experimental results of linear regression for interlaced data as compared to the general linear regression model. The total number of test data is 2,500, and the number of interlaced data is three. So, the proposed approach is evaluated by generating three linear regression models from the interlaced feature data and combining the respective results. In analyzing the result of classification, predictions are evaluated as correct if the linear models can classify similar and dissimilar software correctly. In the evaluation results of the general linear regression model, 2,252 predictions are correct classification among 2,500 test data, so the overall classification accuracy is 90.08%. In the proposed approach to interlacing feature data, 2,356 predictions are correct classification among 2,500 test data, so the classification accuracy is 94.24% that is higher than that of the previous linear regression model. From the experimental results, we confirm that the proposed approach to applying linear regression models to interlaced data is more effective in software classification.

From the experimental results, the proposed approach to interlacing feature data has increased the classification accuracy by 4% compared to the conventional linear regression method. Because the traditional linear regression models produce only one model generated from training data, if some feature values in training data contain noise, the result of the model can produce wrong predictions. The proposed method in this paper controls the features of training data by separating into several training data, so the effect on prediction results are limited only to the local linear regression model. Thus, combining multiple linear regression models from the interlaced data can improve the accuracy of the similar software classification as compared to the conventional linear regression methods.

5 Conclusion

Linear regression is widely used in prediction problems that can be solved by modeling the relationship between input and output data. In this paper, we have proposed an approach to interlacing data in the linear regression to generate multiple regression models from the interlaced data. We have performed experiments to evaluate the proposed approach as compared to the conventional linear regression. The experimental results show that the proposed method can classify similar software more accurately compared to the general linear regression model. The proposed approach is expected to be an effective approach in linear regression to identify similar software. The proposed approach is also expected to be applied in various areas of prediction problems more accurately.

Acknowledgements This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(Ministry of Education) (No. NRF-2017R1D1A1B03034769).

References

1. Murphy KP (2012) Machine learning: a probabilistic perspective. The MIT Press
2. Shalev-Shwartz S, Ben-David S (2014) Understanding machine learning: from theory to algorithms. Cambridge University Press
3. Domingos P (2012) A few useful things to know about machine learning. Commun ACM 55(10):78–87
4. Lim H-i (2019) A linear regression approach to modeling software characteristics for classifying similar software. Proc COMPSAC 2019:942–943
5. Python programming language, https://www.python.org/
6. scikit-learn: Machine Learning in Python, https://scikit-learn.org/stable/index.html
7. ANTLR (ANother Tool for Language Recognition), https://www.antlr.org/
8. The Jakarta-ORO, https://jakarta.apache.org/oro/

Blockchain and Trusted Execution Environment Based Fairness Incentive Mechanism in Crowdsensing

Yihuai Liang, Yan Li, and Byeong-Seok Shin

Abstract We propose a novel scheme for blockchain-based crowdsensing applications by leveraging Trusted Execution Environments to address the issue that malicious requesters trick workers by publishing abnormal crowdsensing requirements, resulting in that requesters can get sensing data of the workers without needing to pay workers during a crowdsensing dealing. Experiments are performed to evaluate the processing time and network payload of our prototype system.

Keywords Crowdsensing · Blockchain · TEE · Fairness dealing

1 Introduction

People who need a specified kind of sensing data, called requesters, can publish their requirements to a crowdsensing [1] application platform, such as Amazon Mechanical Turk. The requirements contain rewards, sensing data specifications and quality standards. People who have the sensing data that meets the requirements, called workers, can sell it to the requesters. Traditional centralized crowdsensing services belong to companies or organizations, which are not totally trustworthy.

A blockchain can act the role of traditional crowdsensing platform service providers. Firstly, blockchain has a nature of anonymity, can evaluate sensing data via chain code and pay workers mandatorily if sensing data meets the requirements. Workers do not concern about getting unfair rewards if the sensing data meets the requirements, while requesters also do not concern about getting low-quality sensing data after paying.

Y. Liang · Y. Li · B.-S. Shin (✉)
Inha University, Incheon, Republic of Korea
e-mail: bsshin@inha.ac.kr

Y. Liang
e-mail: liangyhgood@gmail.com

Y. Li
e-mail: leeyeon@inha.ac.kr

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_5

However, one issue is that requesters can publish malicious smart contract to the blockchain on purpose to make the sensing data fail to pass the quality evaluation, even though the quality is sufficiently good. If the requesters are also miners in the blockchain, they can access the sensing data. Because data in a blockchain is public and transparent. The worker can not get paid because of failure of the data evaluation. It is unfair for workers, and they will lose their interests severely to participate in crowdsensing tasks if the system has such a vulnerability.

To solve the issue mentioned above, we leverage Trusted Execution Environment (TEE) [2] on blockchain-based crowdsensing and propose a novel scheme. However, simple usage of TEE is not enough to solve our issue. Because a secure area of TEE, called enclave, is to guarantee code and data loaded inside to be protected with respect to confidentiality and integrity. But it cannot prevent requesters to put a malicious smart contract of the requirement into an enclave. The key idea of our proposal is that if and only if the sensing data meets the requirements and workers get paid, the requesters and miners can access the data. To achieve that, we execute the smart contract of requirement inside an enclave, along with the sensing data uploaded by workers through Transport Layer Security (TLS) connection. The enclave outputs value 1 if the sensing data meets the requirements and then pay workers rewards. More details are presented later. Our contribution is as follows:

- To leverage TEE to propose a blockchain-based novel crowsensing scheme to solve the issue that requesters trick workers by publishing malicious sensing data requirements to get the data without needing to pay.

2 Related Work

Recently, researchers have focused on applying blockchain to crowdsensing applications in an effort to address these challenges. However, to date, only a few contributions have been made. Wang et al. [3] presented a blockchain-based privacy-preserving incentive mechanism in crowdsensing applications. In this system, a blockchain network plays the role of a trusted third party. It is responsible for evaluating the quality of sensing data and the payment for workers providing the data. The paper also proposes intra-group negotiation and k-anonymity privacy protections. Jia et al. [4] proposed a crowdsensing system focusing on solving the issue of sensing data leaking workers' location privacy when they upload data to the blockchain. Li et al. [5] proposed a blockchain-based decentralized framework for crowdsourcing mainly to solve the problem of a single point of failure in the traditional crowdsourcing system. These valuable contributions solve problems regarding reliability, privacy-preservation, and security. However, the challenge of requesters publishing malicious requirements to the blockchain to trick workers in existing crowdsensing systems has not been discussed. To the best of our knowledge, this is the first paper to focus on addressing this pertinent issue in blockchain-based crowdsensing applications.

Bowman et al. [6] presented private data objects (PDO), which enables a privacy-preserving approach to run smart contracts over private data by leveraging Intel Secure Guard Extensions (SGX) [7, 8]. However, smart contracts are provided by a trusted party, and the user must communicate with the contract maker off-chain to verify and approve the contract before using it. In contrast, our proposal avoids verifying the requirement contract created by requesters but can still guarantee the fairness of crowdsensing transactions.

3 Architecture and Operations of Our Proposed System

A blockchain network can replace the role of traditional centralized crowdsensing services, also provide extra some outstanding utilities. On the one hand, to make a dealing fair for both requesters and workers, we leverage the blockchain network to replace the trusted third party. Such that the requesters and workers need not trust anybody. Also, the blockchain can also provide anonymity to protect user's identities and privacy. Therefore, our architecture, shown in Fig. 1, contains a blockchain network that is the main. On the other hand, the data store in the blockchain is public, which causes the issue we mentioned above. This issue is crucial because it is unfiar for workers. They will lose interests to participate crowdsensing tasks. Therefore, we leverage TEE and well-designed operations and procedures to make blockchain-based crowdsensing applications fair for both the worker and the requester.

Three main parties are involved in the system shown in Fig. 1, which are miners, requesters, and workers. Miners are in charge of collecting, validating transactions and creating blocks for the blockchain. They can gain mining rewards and transaction

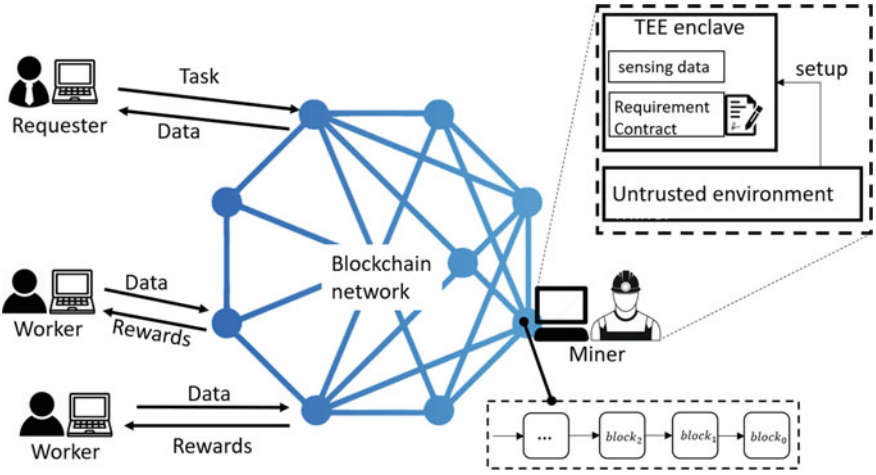


Fig. 1 The architecture of our crowdsensing system based on blockchain and TEE

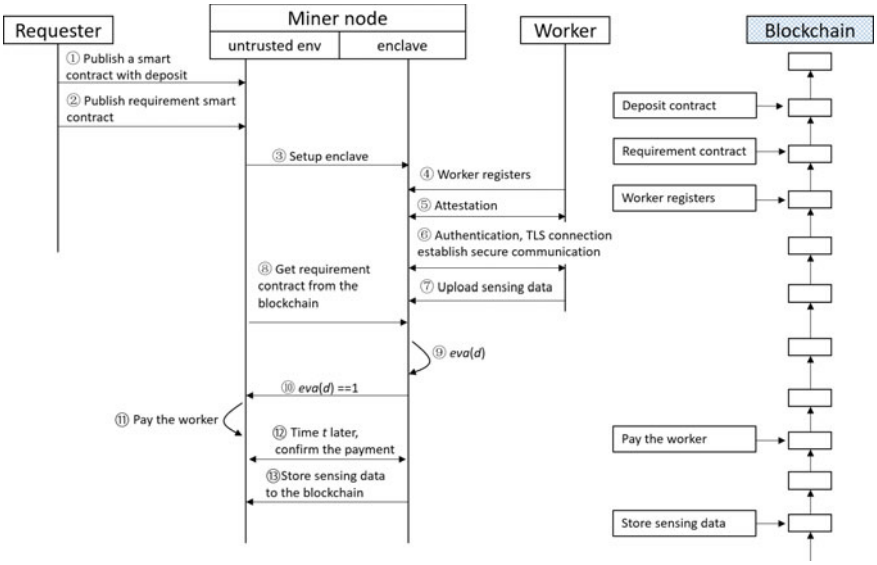


Fig. 2 Operations of our proposed crowdsensing scheme

fees if succeed to create a block. A miner can be a requester to publish tasks or a worker to take tasks. A miner node consists of two parts mainly, an untrusted environment and secure TEE enclaves. Requesters who need sensing data publish tasks and requirements to the blockchain and pay workers who can upload sensing data to the blockchain if the data meets the requirements. Workers are persons who accept the data sensing task published by requesters in the blockchain. They upload sensing data and get rewards.

Figure 2 shows the operations of our proposed crowdsensing scheme. (Steps ①–②) The requester submits a smart contract (i.e. chain code) with a deposit to blockchain before publishing a requirement smart contract. The deposit is to guarantee the requester enables to pay workers if the sensing data meets the requirements. (Steps ③–⑦) Then, a secure enclave is set up by the untrusted environment in the miner node. The worker performs attestation with the enclave residing on the miner node to validate the enclave is set up properly in the miner node. If the attestation passes, the worker establishes secure communication with the enclave through TLS connection and upload sensing data to the enclave. (Steps ⑧–⑪) The enclave concurrently gets the requirement contract published by the requester from the blockchain. The miner node executes the requirement contract with the sensing data inside the enclave to evaluate the quality of the data. If the quality meets the requirements, the enclave outputs value 1 (i.e. $eva(d) == 1$) to the untrusted environment of the miner node. The untrusted environment is in charge of paying rewards to the worker by using the deposit of the requester. Else if $eva(d) == 0$, the enclave deletes the sensing data and terminates the transaction. (Steps ⑫–⑬) After time t, the enclave validates the payment and stores the sensing data in the blockchain. The time t is a

parameter depending on the blockchain security strategy, such as 60 min in Bitcoin. If the payment validation fails, the enclave will delete the data and terminate the transaction. The part on the right of Fig. 2 shows data stored in the blockchain or operations that change the states of the blockchain.

4 Experiment

We leverage Intel SGX as a TEE included in recent Intel CPUs which introduces the concept of isolated hardware enclaves that can be created and managed using new CPU instructions. The experiment setup consists of three machines: An Intel SGX-enabled machine running Ubuntu 18.04 LTS, where we install Ethereum [9] client, OpenEnclave [10] as an SDK for building enclave under Intel SGX hardware, Microsoft eEVM [11] as a standalone Enclave ready Ethereum Virtual Machine running inside an enclave to execute smart contracts. This machine acts as a miner node. In addition, we implement a server in C++ language to receive and response requests from workers and requesters atop Ethereum. This server runs in the miner node and interacts with the Ethereum client via JSON RPC API. TEE enclaves run in our server as security areas. The second and third takes the role of workers and requesters, respectively.

We evaluate the performance of our prototype system by calculating the processing time and network transmission data among a worker, a requester, and a miner, comparing with a naïve and Unsafe Crowdsensing Scheme (UCS) but blockchain-based (Fig. 3). The difference between our proposal and UCS is that UCS does not use TEE, remote attestation or TLS connection. So, as soon as the sensing data is uploaded to the miner node, it becomes public and transparent, which can be accessed by everyone. In addition, the sensing data is stored in the blockchain as soon as $eva(d) == 1$ without needing to wait the time t for payment validation. In Fig. 3, the x-axis is the transaction number. Note that a transaction here is from the time that a requester

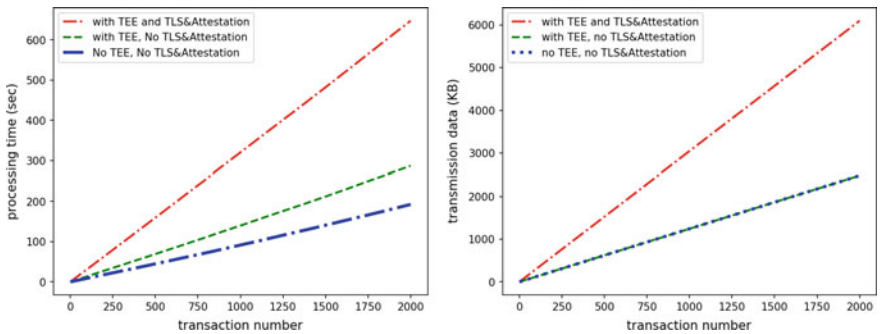


Fig. 3 Processing time (left) and transaction data (left) with transactions

publishes a deposit smart contract to the time that the sensing data is stored in the blockchain, i.e. it includes all steps in Fig. 2.

But on the left of Fig. 3, we only consider the processing time of the miner node, excluding requesters and workers. Also do not include the time of waiting for a transaction validation accepted into a blockchain. We abbreviate Remote Attestation and TLS to RATT. Our proposed system is equipped with TEE and RATT. That's why its processing time is longer than the case of UCS for each transaction. If we remove RATT from our proposed system, the processing time shown by the middle line in the left sub-figure is almost near the line at the bottom. So, we can get a conclusion that the processing of RATT needs much more CPU resource and it brings in the most CPU payload for our proposal comparing with UCS. In the right sub-figure of Fig. 3, we present the network transmission data of our proposed system, shown by the top line. If we remove RATT from the system, the transmission data is the same as UCS in the way that was expected. Because the network communications between workers and miners or between requesters and miners are the same as UCS if our proposal system excludes RATT. So from this sub-figure, we can get a conclusion that RATT brings in all network traffic payload for our proposed system.

5 Analysis and Conclusion

Our proposed scheme overcomes the issue that malicious requesters can get sensing data without paying workers by publishing abnormal requirements that always make $eva(d) == 0$. Firstly, if a requester is not a miner of the blockchain, he can only get the sensing data after the data is stored in the blockchain. In this case, the requester can get the data that meets his requirements but get nothing if the data does not. The worker does not need to concern about being tricked by the requester by making a malicious requirement contract. Secondly, if the requester is also a miner of the blockchain, TEE guarantees that the miner can only access the sensing data if $eva(d) == 1$ and the workers get paid.

Acknowledgement This research was supported by Research Program To Solve Social Issues of the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT. (No. NRF-2019R1A2C1090713).

References

1. Crowdsensing (2020). https://en.wikipedia.org/wiki/Crowdsensing
2. Sabt M et al (2015) Trusted execution environment: what it is, and what it is not. In: 2015 IEEE Trustcom/BigDataSE/ISPA, vol 1. IEEE
3. Wang J et al (2018) A blockchain based privacy-preserving incentive mechanism in crowdsensing applications. IEEE Access 6:17545–17556

4. Jia B et al (2018) A blockchain-based location privacy protection incentive mechanism in crowd sensing networks. Sensors 18(11):3894
5. Li M et al (2018) Crowdbc: a blockchain-based decentralized framework for crowdsourcing. IEEE Trans Parallel Distrib Syst 30(6):1251–1266
6. Bowman M, Miele A, Steiner M, Vavala B (2018) Private data objects: an overview. http://arxiv.org/abs/1807.05686
7. Costan V, Devadas S (2016) Intel SGX explained. IACR Cryptology ePrint Arch 2016(086):1–118
8. McKeen F et al (2013) Innovative instructions and software model for isolated execution. Hasp@ isca 10(1)
9. Buterin V (2014) Ethereum white paper: a next generation smart contract & decentralized application platform. First version 53
10. Openenclave (2020). https://github.com/openenclave/openenclave
11. Microsoft eEVM (2020). https://github.com/microsoft/eEVM

Identifying Critical Topics for Successful Games in Game Reviews by Applying Latent Dirichlet Allocation

Mookyung Kwak, Ji Su Park, and Jin Gon Shon

Abstract What makes successful games? Answering this question would be beneficial not only for game industry but also for various other fields like education in which gamification becomes important technique. In this study, we extract critical topics from the game reviews using topic modeling approach, Latent Dirichlet Allocation, then model general principles behind the topics in which we will discuss the validity and usefulness for gamification.

Keywords Topic modeling · Game reviews · Latent dirichlet allocation

1 Introduction

Gamification becomes universal phenomenon around the fields widely. We have various frameworks which describe how the gamification techniques can be applied. However, we still suffer lack of empirical and extensible resources that can give us practical implications. Only implications which based on substantive material can stand test of time. When gamification techniques need to be applied to real world applications, they are bumped against so many obstacles that most of them become useless. At any rate, educators or entrepreneurs in the field need to find practical and feasible solutions which may apply to various conditions. If we have implications which have strong roots in reality, we can consider extensible use cases by adding flexible interpretations.

It can be supposed that game reviews, as a whole set of words, sit on latent topics which will show critical factors to become successful games. In that sense, we will

M. Kwak · J. G. Shon (✉)
Department of e-Learning, Korea National Open University, Seoul, Korea
e-mail: jgshon@knou.ac.kr

M. Kwak
e-mail: pporri11@knou.ac.kr

J. S. Park
Department of Computer Science and Engineering, Jeonju University, Jeonju, Korea
e-mail: jisupark@jj.ac.kr

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_6

find topics that lie beneath words in game reviews. We define topics here as the subject that reviewers are trying to say in their reviews. Therefore, the latent topics we find in game reviews, presumably hold meaningful solutions that can help games more successful and gold standard that differentiate successful ones.

In this study, we apply Latent Dirichlet Allocation (LDA) technique to reveal the latent topics and model them. The technique will identify for us combined words from game reviews on which we should define topics. With the topics, we will discuss which one can be used for gamification in the following sections.

2 Background

Amory et al. [1] endeavor at presenting a model, called "Game Object Model", which integrates educational theory and game design. This model is based on constructivist educational theory where the importance of building or constructing knowledge by learners themselves is emphasized. Amory et al. also emphasized the importance of models which provide developers with a conceptual and practical framework that can support the development process.

King et al. proposed a new taxonomy of video game features in [5], which includes: (a) social features, (b) manipulation and control features, (c) narrative and identity features, (d) reward and punishment features, and (e) presentation features. This approach is out of rigorous attempt to classify and organize the psycho-structural elements of video games.

Said focused on what makes users engaged in the educational games in [6]. If all the designing features like simulation interaction, construct interaction, immediacy, feedback, goals are present, the game could achieve high engagement level, he believed.

All these models and frameworks are focusing on what components or elements are essential for game design, that is surely important. Nevertheless, game designers in field require practical techniques which can be adopted in lower level. For example, to add reward and punishment features, designers should define what could be reward and punishment in the game system. And then they should consider what is trending nowadays to make the features more appealing to users. In these processes, the need for having a guide which has concrete and rich language but also has room for interpretations that can give better adaptation opportunities. We thought game reviews are concealing data enough for making this guide possible. This leads naturally to that extracting meaningful factors from game reviews and it becomes main concern of our study.

3 Topic Modeling Analysis

3.1 Theoretical Framework

To extract meaningful factors for game, we used Topic Modeling Analysis in game reviews. It is supposed that game reviews include topics which are talking about success factors in game, and the topics lie dormant in words in the reviews. Topic Modeling Analysis can reveal the topics in the form of enumerated words which need to be inferred by researchers.

In the accuracy manner of extracting meaningful factors in reviews, humans can outperform any algorithms, but using humans for this task has two problems: first, processing large volume of texts is costly, error-prone work. Secondly, the decisions can be arbitrary, which makes reproducing the result hard.

Latent Dirichlet Allocation (LDA) is one of Topic Modeling Analysis technique. Blei et al. described this as a generative probabilistic model for collections of discrete data such as text corpora in [2]. At any rate, the main reason for choosing in this study is that its results are fairly understandable and interpretable by human intelligence.

The essence of methodologies in the field of information retrieval is how to convert text corpora into numbers and reduce the dimensions of those numbers, which is called dimensionality reduction.

The first subject, converting text corpora into numbers, is including frequency count of words in document, the order of words in document and probability of word distribution. For example, simple bag of words model only reflects frequency count of words in documents, meanwhile Latent Semantic Indexing (LSI) arguably considers semantic position of each words.

Probabilistic LSI emerged in this context. Hoffman proposed on [4] that each document can be represented as proportional mixtures of fixed set of topics which also mean that dimensionality reduction can be applied to produce the probability distribution of the topics. This idea provided important step to LDA methodology. LDA adopted this probabilistic modeling to the level of documents.

The basic idea is that documents are represented as random mixtures over latent topics, where each topic is characterized by a distribution over words.

In this modeling, the topics are latent beneath documents and words, so the result of analysis shows only a list of probabilistically strongly connected words which need to be interpreted to topics by researchers' own words.

3.2 Methodology

Firstly, the game reviews are collected from the game review site (www.metacritic.com/game) by web-site crawling technique. The reviews date between 1996 and 2019. Crawled reviews are 8,465 due to dead or changed links. After cleaning errors and duplicated reviews, 7,745 reviews survived.

Table 1 Examples of 50 topics with supporting top 10 words

Number	Words
0	enemi level use attack time like boss make power play
1	charact play array releas like make fighter version level adventur
2	wii version control ps array consol remot like play year
	…
49	ds mario nintendo mini super screen stylus touch titl play

A corpus is made after natural language processing technique being applied. All words are lowercased, and punctuations are removed. As mentioned in Sect. 3.1, Latent Dirichlet Allocation is used to extract topics from the corpus.

To decide proper hyper-parameter in LDA analysis, perplexity number is used which showed 50 topics would be adequate for this analysis. The result is shown in Table 1.

Words are stemmed and lemmatized to be identified as their root forms. After LDA is performed, it is supposed to infer topics from the words by researcher's intuition and experience. Table 2 shows the whole set of inferred topics.

4 Discussions

4.1 Stimulation—Response—Desire—Goal (SRDG) Model

Categorizing the topics inferred above is one of possible way to obtain implications. After looking through all the topics one by one, we concluded that all the topics are aimed mainly at maintaining attention from users by all means. Corbetta and Shulman [3] proposed that attentional functions are dependent on two types of systems—goal-directed (top-down) and stimulus driven (bottom-up)—which are from totally different part of brain areas. Corbetta and Shulman provided neurophysiological evidences that primates inherently use these different strategies to process information efficiently.

We adapt this approach to the model developed from inferred topics. Game developers use various stimuli to catch user's attention, but proper response is also needed to keep the attention. This is bottom-up approach. But gamers usually expect more than just sensory stimulus, thus game developers also provide various objects of

Table 2 50 inferred topics from words

Topics		
Stacking up level	Being a hero	Medieval times fantasy
Adventure	Outrageous power	Space fantasy
Sense of physical control	Flash visual stimulation	Odd experience
Hunting prey	Being social through party game	Learning process
Storytelling	Achieving mission	Creative play
Time attack	Giving the entire picture	Fast reflection
Feel the rhythm of music	Easy but continuous challenge	My own little world
Build up	Adventure in teamwork	Experience real world
Fan of franchise	Imitating sport play	Explore unrealistic world
Belching violence	Planning strategy	Sensation of speed
Eye-hand coordination	Frantic stimulus–response cycle	Becoming a super-power
Crush everything except me	Getting richer	Solving puzzle
Simulating reality elaborately	Defeat others	Overcome invincible adversity
Ear-hand coordination	In the shoes of a character	Bragging intelligence
Special skills	Fan of character	Spectacular scenes
Watch others to play	Using items creatively	Old franchise, new game
Improving skills	Planning everything	

desires or goals which can direct users to select specific stimulus and keep the attention. This can be regarded as top-down approach.

In this model, we introduced four major categories which are stimulation, response, desire, and goal. The first two categories are related to bottom-up approach, whilst the last two are top-down ones. Table 3 shows categories and sub-categories that group the topics.

Table 3 Topics categorized into SRDG model

Category	Sub-category	Topics
Stimulation	Low-level stimulus	Flash visual stimulation
		Ear-hand coordination
		Sensation of speed
	High-level stimulus	Fan of franchise
		Storytelling
		Giving the entire picture
		Imitating sport play
		Fan of character
		Medieval times fantasy
		Space fantasy
		Old franchise, new game
Response	Impromptu response	Time attack
		Feel the rhythm of music
		Eye-hand coordination
		Sense of physical control
		Ear-hand coordination
		Frantic stimulus–response cycle
		Fast reflection
	Strategic response	Build up
		Planning strategy
		Using items creatively
		Planning everything
		Creative play
Desire	Desire of defeating	Belching violence
		Crush everything except me
		Watch others to play
		Special skills
		Outrageous power
		Defeat others
	Desire of achieving	Hunting prey

(continued)

Table 3 (continued)

Category	Sub-category	Topics
		Improving skills
		Stacking up level
		Achieving mission
		Easy but continuous challenge
		Getting richer
		Solving puzzle
	Desire of learning	Simulating reality elaborately
		Learning process
		Experience real world
	Desire of being social	Being social through party game
		Adventure in teamwork
	Desire of new world	Adventure
		Odd experience
		Explore unrealistic world
		Spectacular scenes
	Desire of being others	Being a hero
		In the shoes of a character
		Becoming a super-hero
Goal	Internal growth	Overcome invincible adversity
	External showing off	Bragging intelligence

4.2 Application and Development of the Model

We think the value of this model resides in sub-categories. These sub-categories are not deterministic, they can be extended as new technologies and trends emerge. But developers still can get hints of what kind of strategies can be applied when they make games.

When they design games, the categories and sub-categories are working as good guidelines to depend on. For example, developers can evaluate whether their game has sufficient features for both bottom-up and top-down approaches. And if they think bottom-up or top-down features are not enough, more detailed and practical strategies can be found by looking through the sub-categories and topics.

This model also can be applied to another field like education. Not surprisingly, not all the topics are applicable to the field, though game developers strain every nerve to fill up gaps between players' attention. Educators can make similar efforts. For example, they can re-design curriculum to apply 'desire of being others' or 'high-level stimulus', as it were, let students make decisions themselves on the basis of historical facts or tell students stories behind mathematical formulae.

5 Conclusion

We focused on extracting useful implications from big data which can be reproduced by any following study. We believe these implications could be more useful when conventional psychological and neurophysiological knowledge apply to the implementation as like as being applied in making SRDG model. We hope these topics and model contribute to reveal how games can be so addictive. If we can develop more robust and concrete theory on the reason of addiction, it will lead to using of them in controlled ways, so that it would not be impossible task to take advantage of gamification in more frequent and deeper occasions.

References

1. Amory A, Seagram R (2003) Educational game models: conceptualization and evaluation: the practice of higher education. South Afr J Higher Educ 17(2):206–217
2. Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. J Mach Learn Res 3:993–1022
3. Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3(3):201
4. Hofmann T (1999) Probabilistic latent semantic analysis. In: Proceedings of the fifteenth conference on uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc., pp 289–296
5. King D, Delfabbro P, Griffiths M (2010) Video game structural characteristics: a new psychological taxonomy. Int J Mental Health Addict 8(1):90–106
6. Said NS (2004) An engaging multimedia design model. In: Proceedings of the 2004 conference on Interaction design and children: building a community, pp 169–172

Design of a Multi-core IP-NDN Gateway for Smart Dust IoT Environments

JuGeon Pak, Joonsuu Park, and KeeHyun Park

Abstract In a smart dust IoT (Internet of Things) environment, thousands of sensors are deployed and hence a great number of sensed data is to be processed. In this paper, an IP-NDN (Named Data Networking) gateway is proposed in order to enhance data processing performance in a smart dust IoT environment. An IP-NDN gateway provides links between IP networks and named data networks. The gateway is designed to allocate cores in an IP side and a NDN side separately in such a way that it provides location independence and hence achieves enhanced data processing performance.

Keywords Smart dust · IoT · IP network · NDN network · IP-NDN gateway · Multi-core

1 Introduction

A smart dust IoT (Internet of Things) system is a kind of IoT [1–3] systems where a lot of very small sized sensors are scattered from the air to gather information on rough terrain such as the Amazon rainforest or the moon's surface that is difficult for humans to access [4–6]. In a smart dust IoT environment, thousands of sensors are deployed and hence a great number of sensed data is to be processed. However, network performance in these extreme environments is very limited.

In this paper, an IP-NDN (Internet Protocol-Named Data Networking) gateway is proposed in order to enhance data processing performance in a smart dust IoT environment. The NDN is an architecture that communicate not by an IP address but by data name itself and it is considered to be one of appropriate architectures in IoT

J. Pak · J. Park · K. Park (✉)
Department of Computer Engineering, Keimyung University, Daegu, Republic of Korea
e-mail: khp@kmu.ac.kr

J. Pak
e-mail: jugeonpak@gmail.com

J. Park
e-mail: parkjoonsuu@gmail.com

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_7

environments where a large number of data is processed [7–9]. However, it is very difficult to apply the NDN policies to all the network nodes at once.

Therefore, an IP-NDN gateway is considered to be one of solutions [10, 11]. An IP-NDN gateway provides links between IP networks and NDN networks. The gateway in this study is designed to allocate cores in an IP side and a NDN side separately in such a way that it provides location independence and hence achieves enhanced data processing performance. In this paper, we propose a system architecture that improves data processing performance by integrating an NDN-friendly Smart Dust IoT environment and provides an IP-friendly communication network through an IP-NDN gateway.

The remainder of this paper is organized as follows: Sect. 2 introduces a smart dust IoT system and Sect. 3 proposes an IP-NDN gateway for smart dust IoT environments. Conclusion and future studies are discussed in Sect. 4.

2 Smart Dust IoT System

Most of the IoT environments use IP networks [12, 13]. IoT environments that cover large areas such as smart cities also use IP networks [14]. This is because, in general IoT environments, IP networks are used as infrastructure networks [13]. Using IP networks in IoT environments may not be a big issue, but, in smart dust IoT environments where a large amount of devices transmits data, it may cause a bottleneck [15]. NDN can alleviate bottleneck problems and support data verification by introducing a content store (intermediate node that stores content) [16].

As depicted in Fig. 1, a smart dust IoT system is an IoT systems where a lot of very small sized smart sensors are scattered from the air to gather information on rough terrain such as the Amazon rainforest or the moon's surface that is difficult for humans to access. Aircrafts or balloons can be used to distribute the sensors in

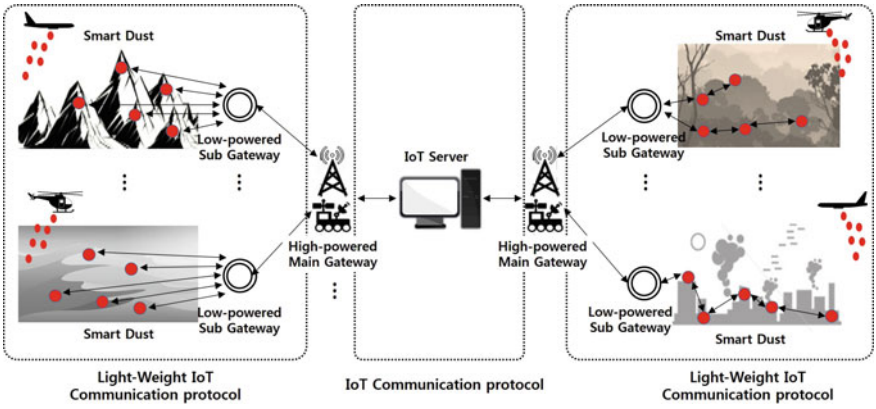


Fig. 1 Overview of a smart dust system

the air. Once landed on the ground, the smart dust sensors begin to sense data in the surrounding area in order to send the data to low-powered sub gateways which are scattered from the air also. Then, the low-powered sub gateways receive the data and perform some processing to send it to high-powered main gateways. The high-powered main gateways can be stationary or mobile. The high-powered main gateways communicate with the IoT server via IoT communication protocols using IP networks or NDN networks.

3 IP-NDN Gateway

Located between IP networks and NDN networks, the IP-NDN gateway proposed in this study provides the interconnection between the two networks. Figure 2 shows the structure of the IP-NDN gateway proposed in this study. The proposed gateway consists of the NDN side, the IP side, and the Shared Memory. The IP-NDN gateway is designed to work in multi-core environments. The NDN side and the IP side are designed to work in separate cores, which guarantees location dependence and enhanced packet processing performance as well.

In the NDN side, there exist the Packet handler and the Routing/Forwarding Engine. The Packet Handler performs packet encapsulation/decapsulation, tunneling, and conversion. Data in the two networks are relayed by the Routing/Forwarding Engine. The tables in the Shared Memory are updated according to the analysis of transmitted packet packets in the NDN network. In the IP side, the Routing SW analyzes routing packets transmitted in the IP network to update the tables in the Shared Memory.

The Shared Memory consists of the tables to be referenced for data routing/forwarding between an IP network and a NDN network. The Forwarding Information Base (FIB) table stores the mapping relations between data names and the interface list which the data is forwarded to. The Pending Interest Table (PIT) stores the mapping relations between data names and the interface which receives the request of the data. The Content Store (CS) stores the actual data (These tables were

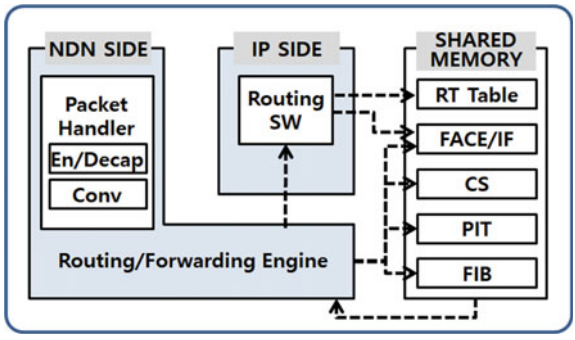


Fig. 2 Structure of the proposed IP-NDN gateway

mentioned in [7]). These tables are references when data are relayed to the NDN side. The Routing Table (RT) is updated by the IP side and referenced when data are relayed to the IP side. Finally, the FACE/IF table stores the physical interfaces which data packet is routed/forwarded to.

4 Conclusion and Future Studies

In this study, an IP-NDN gateway is proposed and designed to provide a bridge between IP networks and NDN networks in smart dust IoT environments. The IP-NDN gateway proposed in this study is designed to work in multi-core environments. A core is assigned to the IP side and the NDN side separately. The gateway processes protocol conversion procedures and packet forwarding procedures independently in separate sides in order to enhance network performance. Validation and implementation works of the design will be performed in the near future.

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 SR (2017) Internet of Things: architectures, protocols, and applications. J Electr Comput Eng 2017
2. Mao X, Li K, Zhang Z, Liang J (2017) Design and implementation of a new smart home control system based on internet of things. In: 2017 International Smart Cities Conference (ISC2). IEEE, pp 1–5
3. Rahmani AM, Gia TN, Negash B, Anzanpour A, Azimi I, Jiang M, Liljeberg P (2018) Exploiting smart e-Health gateways at the edge of healthcare Internet-of-Things: a fog computing approach. Future Gener Comput Syst 78:641–658
4. Kahn JM, Katz RH, Pister KSJ (1999) Next century challenges: mobile networking for "Smart Dust". Paper presented at the Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking, Seattle, Washington, USA
5. Warneke B, Last M, Liebowitz B, Pister KS (2001) Smart dust: communicating with a cubic-millimeter computer. Computer 34(1):44–51
6. Park J, Park K (2019) A dynamic plane partition method for DPDK in smart dust environment. In: Proceedings of 2019 IEEE 10th annual information technology, electronics and mobile communication convergence (IEMCON2019). Vancouver, pp 562–567
7. Zhang L, Afanasyev A, Burke J, Jacobson V, Claffy K, Crowley P, Papadopoulos C, Wang L, Zhang B (2014) Named data networking. ACM SIGCOMM Comput Commun Rev 66–73
8. Amadeo M, Campolo C, Iera A, Molinaro A (2014) Named data networking for IoT: an architectural perspective. In: 2014 European conference on networks and communications, pp 1–5
9. Arshad S, Azam MA, Rehmani MH, Loo J (2019) Recent advances in Information-Centric Networking-based Internet of Things (ICN-IoT). IEEE Internet of Things J 6(2):2128–2158

10. Refaei T, Ma J, Ha S (2017) Integrating IP and NDN through an extensible IP-NDN gateway. In: Proceedings of the 4th ACM conference on information-centric networking, pp 224–225
11. Quevedo J, Ferreira R, Guimaraes C, Aguiar RL, Corujo D (2017) Internet of Things discovery in interoperable information centric and IP networks. Internet of Things discovery in interoperable information centric and IP networks. Internet Technol Lett 1(1):1–6
12. de Almeida IBF, Mendes LL, Rodrigues JJ, da Cruz MA (2019) 5G Waveforms for IoT applications. IEEE Commun Surv Tutorials 21(3):2554–2567
13. Yaqoob I, Ahmed E, Hashem IAT, Ahmed AIA, Gani A, Guizani M (2017) Internet of Things architecture: recent advances, taxonomy, requirements, and open challenges. IEEE Wirel Commun 24(3):10–16
14. Gubbi J, Buyya R, Marusic S, Palaniswami M (2013) Internet of Things (IoT): a vision, architectural elements, and future directions. Future Gener Comput Syst 29(7):1645–1660
15. Park J, Park K (2020) A dynamic plane prediction method using the extended frame in smart dust IoT environments. Sensors 20(5)
16. Hail MA, Amadeo M, Molinaro A, Fischer S (2015) Caching in named data networking for the wireless internet of things. In: 2015 international conference on recent advances in Internet of Things (RIoT), pp 1–6

A Study on Protocol Comparison for Energy-Efficient Network Configuration in Mobile Edge Computing

Bok Gi Min, Ji Su Park, and Jin Gon Shon

Abstract Among the Internet of Things (IoTs) devices, when mobile user equipments (MUEs) such as smart phones or laptops that utilize direct Internet services increase, there is a problem of service delay due to network load when sending data to directly connecting. In order to solve this, Mobile Edge Computing (MEC), which is trying to solve the service delay problem by providing services in the network edge area, which is an intermediate area between Internet services and MUEs, is attracting attention. Among the many limitations of MEC, due to the characteristics of battery-powered MUEs, there is a need for an energy-efficient wireless network configuration in an energy-limited environment. A variety of studies have been conducted in the field of Wireless Sensor Networks (WSNs) to configure multiple nodes and energy-efficient wireless networks in an energy-limited environment, as in the MEC environment. In this paper, we propose a method of applying cluster-based LEACH (low energy adaptive clustering hierarchy) and chain-based PEGASIS (Power-Efficient Gathering in Sensor Information Systems) to MEC environments among typical WSN routing protocols. Simulation results show that the proposed network configuration can increase the network life compared to the existing configuration that directly connects to the MEC.

Keywords Mobile edge computing · LEACH · PEGASIS · Energy efficient · Wireless sensor networks · Network life time

B. G. Min · J. G. Shon (✉)
Department of Computer Science, Graduate School, Korea National Open University, Seoul, Korea
e-mail: jgshon@knou.ac.kr

B. G. Min
e-mail: minbknim@gmail.com

J. S. Park
Department of Computer Science, Jeonju University, Jeonju, Korea
e-mail: jisupark@jj.ac.kr

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_8

1 Introduction

In IoT environment, various sensors or devices (or things) can connect to internet services over a network. As the number of high-performance MUEs among smartphones, laptops and tablets among IoT devices has increased, large-scale processing requests are possible in the short term. However, the network's physical limitations led to bottlenecks and service delays [1].

Edge computing has been proposed to reduce network load by providing services directly at the network edge or by processing data and selecting data to send to Internet services. In addition, as the number of MUEs increases, research on MECs specialized for mobile environments among edge computing is underway [2, 3]. Network life time is important because MEC is a wireless network environment that cannot provide continuous energy. In WSN has been studied for a long time to extend network life through network configuration.

In this paper proposes a method of applying cluster-based LEACH and chain-based PEGASIS, which are representative network configurations studied in WSN, to MEC environments. In addition, the proposed network configuration is compared with the MEC environment through simulation to find a suitable network configuration.

2 Related Work

2.1 Mobile Edge Computing

Due to the slow development of battery technology compared to processor development, MUE has many limitations on the use of Internet services. To address this, MCC (Mobile Cloud Computing) was developed. MCC is a method of using MUE to use powerful long-distance cloud computing and storage resources over a mobile operator or the Internet's core network (CN) [4]. However, with MCC, there is an enormous additional load on the network, and the network topology delays the service, requiring data to be sent to servers far away from the user. To address the service delay issue, MEC has emerged as a way to provide cloud services to users. The advantage of MCC is that it maintains and eliminates service delays by providing the cloud services needed at the network edge close to the user.

2.2 Network Configuration of WSNs

Network configuration of WSNs consists of topologies and routing protocols. Because the topology is wireless, it follows a logical topology that shows the actual data flow of links and nodes rather than a physical topology. Most of the research was derived mainly from cluster-based topology LEACH [5]. In addition, research

is underway on PEGASIS [6], chain-based topology architecture that improves cluster-based topologies. The routing protocol is divided into data-oriented routing protocol, location-based routing protocol, and hierarchical routing protocol. LEACH and PEGASIS use a hierarchical routing protocol that reduces unnecessary energy waste due to redundant transmission of similar data from neighboring nodes and reduces the load on relay nodes.

3 A Comparative Study

3.1 Problem Statement

The differences between WSN and MEC with the network configuration shown in Fig. 1 are as follows:

(1) WSN's Sensor Node (SN) has the same configuration, but in the MEC environment, various IoTs and MEUs exist.
(2) The SN of the WSN only needs to transmit the data of the sensor, but as various nodes exist, various data are exchanged.
(3) WSN's Cluster Header (CH) produces and transmits sensor data with the same configuration as SN. However, there are various MEUs in MEC and relay data specified as an Access Point (AP).
(4) The role of WSN's base station (BS) is responsible for network configuration and data collection. However, the MEC should be able to connect MEUs to the Internet CN as an AP and also provide some services of the MCC in the edge area.

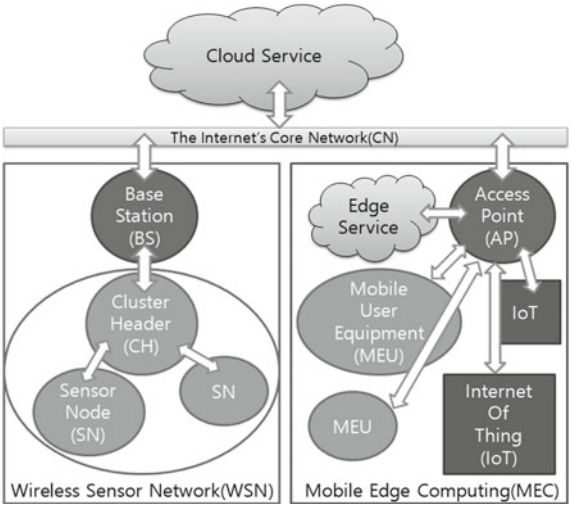


Fig. 1 WSN and MEC network configuration comparison

3.2 Study Condition

This paper assumes the following network environment to refer to the network configuration of WSN. MEUs use batteries and wireless networks with the same performance. MEUs are randomly distributed in a two-dimensional plane. AP and MEC are not separated, but are determined by one AP and have a certain distance from MEUs. MEUs periodically send and receive data through the AP. Every MEUs has a unique identification code. AP is not limited by energy because it can receive energy continuously. All MEUs can know the remaining energy and hourly consumption. All MEUs know the AP's signal strength. All MEUs know the signal strength of the surrounding MEUs.

3.3 LEACH for MEC

LEACH for MEC network configuration, there are no AP services that make up the initial cluster. MCC checks the most used services during T-hours through each AP in the MEU. The services whose priority is confirmed in the MCC are distributed to each AP without overlapping in the local network. The MEU looks for services in the AP of the first connected cluster. If there are no services in the cluster, search on the local network. If you don't have a local network, call MCC for service.

3.4 PEGASIS for MEC

PEGASIS for MEC network configuration does not have AP services that make up the initial chain. Configure the chain network according to the distance from each AP to the local network core. MCC identifies the most frequently used services during T-hours through each AP in the MEU. MCC analyzes the requested service for each chain and distributes the service to the AP without overlapping the chain. MEU searches for services on the first connected AP. If there is no service on the connected AP, the service is found in the chain. And if there is no service even on the local network, ask MCC for service.

4 Performance Evaluation

4.1 Simulation Environment

In order to compare the lifespan of MEUs according to the network configuration, simulations are performed under the same conditions. A local network of 100 MEUs

Table 1 Parameters of simulations

Parameter	Value
Initial energy of MUE	0.5 J
Transmission/reception consumption energy	50 nJ/bit
Free path signal amplification energy	10 pJ/bit/m^2
Multipath signal amplified energy	0.0013 pJ/bit/m^2
Data aggregation energy	5 nJ/bit
Transfer data size (Request/Response)	4000 bits

randomly arranged in a two-dimensional space of 100 × 100 m. A total of 25 APs are placed in each of 20 × 20 m, and the CN and local network are wired. The service provided by MCC or MEC is limited to one HTTP service that responds once to one request. MEU requests service 10 times per second. MCC is scale-out, so there is no limitation, but one HTTP service of MEC can respond 20 times a second. As an environment variable for simulation, MEC conditions are added to the experimental environment of a general wireless sensor network as shown in Table 1. To evaluate the performance of a network configuration, NS2 (Network Simulator Version 2) is used to compare it to other configurations.

4.2 Simulation Result

Figure 2 shows an average of 100 simulations to compare the network lifetimes of MEC, LEACH for MEC and PEGASIS for MEC. The values in Fig. 2 are the average number of transmissions and the average response time when the last MEU dies. For example, the number 3288 in the first row MEC means that the battery has been exhausted since the last MEU averaged 3,288 service requests. As a result, chain-based topology for MEC PEGASIS is 3.2% longer than the lifetime of a typical MEC network configuration, and cluster-based topology for MEC LEACH is 5.6% longer. In addition, when the response time of MEC is 1, the response time of LEACH for MEC is about 3% faster on average.

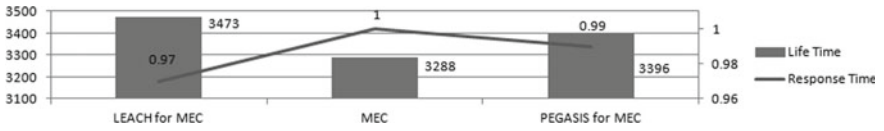


Fig. 2 Network configuration comparison result

5 Conclusion

In This paper proposed PEGASIS for MEC and LEACH for MEC, focusing on the network configuration of WSN. Simulations were performed to compare the effectiveness of the proposed method. Simulation results show that both networks can extend network lifespan compared to typical MECs. Finally, we have the following conclusions:

- Unlike the WSN, in the MEC environment, the network life of the cluster-based topology is longer.
- In the network configuration, the difference in network life is within 10% and the response time is within 3%.
- Additional studies are needed to compare services by diversifying services or diversifying the message length of MUE to make it more realistic.

References

1. Ahmed A, Ahmed E (2016) A survey on mobile edge computing. In: IEEE international conference on intelligent systems and control, pp 1–8
2. Khan WZ et al (2019) Edge computing: a survey. Future Gener Comput Syst 97:219–235
3. Dinh HT, Lee C, Niyato D, Wang P (2013) A survey of mobile cloud computing: architecture, applications, and approaches. Wirel Commun Mob Comput 13:1587–1611
4. Barbarossa S, Sardellitti S, Di Lorenzo P (2014) Communicating while computing: distributed mobile cloud computing over 5G heterogeneous networks. IEEE Signal Process Mag 31(6):45–55
5. Vermareddy SR (2019) Modification of LEACH protocol for improving the stability period and lifetime of WSN. Ph.D. Thesis. Asian Institute of Technology
6. Huarui WU et al (2019) Energy efficient chain based routing protocol for orchard wireless sensor network. J Electr Eng Technol 14(5):2137–2146

A Benchmark Test for Stateless Stream Partitioning Over Distributed Network Environments

Siwoon Son and Yang-Sae Moon

Abstract Distributed stream processing engines (DSPEs) provides various *stateless stream partitioning* to select the receiver tasks for each message regardless of the data fields. A representative DSPE, Apache Storm, provides the polarized stateless stream partitioning: *Shuffle grouping* considering the fairness only and *Local-or-Shuffle grouping* considering the locality only. The recently proposed *Locality Aware grouping* is a novel technique to solve this polarization. However, it is difficult to select an appropriate stream partitioning method considering various configurations of distributed stream applications, network capacity, and data size. In this paper, we benchmark the stateless stream partitioning methods from the perspective of different network bandwidths. To change bandwidths, we experiment on the most widely used the usual Ethernet equipment and the recent InfiniBand, a high-performance network equipment. We can use the benchmark results as the selection criteria for choosing the appropriate stream partitioning method according to the network bandwidth.

Keywords Distributed processing · Real-time processing · Data stream · Locality · Load balancing · InfiniBand

1 Introduction

Distributed stream processing engines (DSPEs) process data streams in real-time in a distributed environment. They provide various *stateless stream partitioning* methods to allow to send to receivers running on the multiple servers regardless of the data fields. As a representative DSPE, Apache Storm [1, 2] provides *Shuffle grouping* (SG) to select receivers in a round-robin manner. This method fairly sends messages to all receivers, but it may cause frequent network communications. Storm also provides *Local-or-Shuffle grouping* (LoSG) to select receivers running in the same process

S. Son · Y.-S. Moon (✉)
Department of Computer Science, Kangwon National University, Chuncheon, South Korea
e-mail: ysmoon@kangwon.ac.kr

S. Son
e-mail: ssw5176@kangwon.ac.kr

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_9

as the sender by considering the locality. However, LoSG may cause a load imbalance problem. To alleviate this polarization of SG and LoSG, the recently proposed *Locality Aware grouping* (LAG) [3] sends more messages to nearby receivers.

In this paper, our main contribution is to benchmark three stateless stream partitioning methods in different network bandwidths. Recently, InfiniBand has been widely used for high bandwidth and low latency among distributed servers [4, 5]. Thus, we measure the performance of Storm grouping methods by changing the bandwidth of InfiniBand as well as Ethernet. The comparative measures are (1) an average transmission time of data streams and (2) a load balancing between tasks. In addition, for various parallel configurations, we assume the cases of sender and receiver balance, sender imbalance, and receiver imbalance.

Extensive experimental results show that SG is the lower the network bandwidth, the higher the message transmission time, and the higher the network bandwidth, the more similar the LAG. LoSG achieves the best transmission time but does not perform load balancing fairly for imbalance cases. Finally, LAG is the lower network bandwidth, the lower transmission time, and the higher network bandwidth, the better load balancing. Through the extensive evaluation, we can develop an adaptable strategy to select an appropriate stream partitioning method based on the network bandwidth.

2 Related Work

Apache Storm [1, 2] is a representative DSPE that quickly processes data streams generated in real-time through distributed servers. Compared to other data stream processing technologies, EsperTech Esper [6] does not provide scalability, and Apache Spark Streaming [7] has a higher latency than Storm [8]. Thus, Storm is a computing framework that is well suited for real-time distributed processing of data streams [9, 10]. An important advantage of Apache Storm is parallelism. In other words, as shown in Fig. 1, Storm replicates tasks to multiple servers to process streams in parallel.

DSPE requires stream partitioning to select receivers to send messages to several receiver tasks. Among the stream partitioning provided by Storm, SG and LoSG are stateless stream partitioning that sends messages regardless of the data fields. SG has the advantage of properly balancing the load on the receivers because the receiver tasks are randomly selected. But SG results in high network latency. Storm also provides LoSG to solve this problem by considering the locality. However, LoSG may cause a load imbalance problem that certain tasks do not work or require much more work than other tasks. We have recently proposed LAG [3] that satisfies the requirements of locality and load balancing. In short, LAG calculates the network distance of the nodes and then sends more messages to the nearby node.

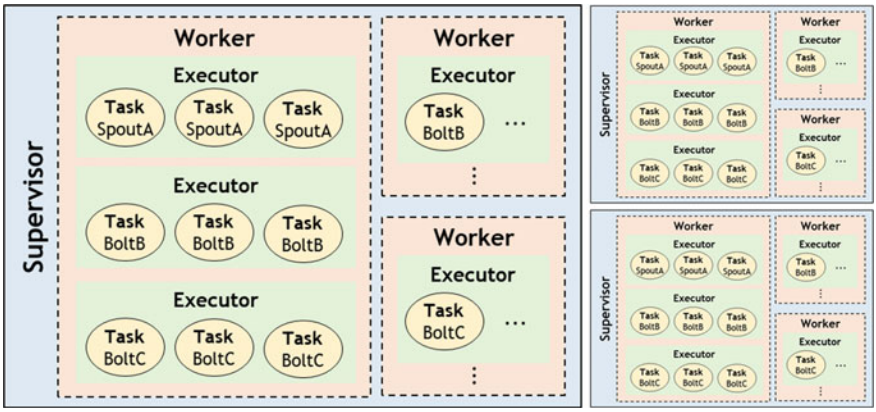


Fig. 1 Parallelism architecture of Apache Storm

3 Design of Benchmark Tests

3.1 Comparative Measures

The first comparative measure is the average transmission time (*ATT*) of stream messages in a job. To measure the ATT, we average the difference between the sent time recorded by the sender task and the received time of the receiver task. The sender tasks send messages at intervals of 100 ms for each message to mitigate performance errors due to network bottlenecks, and we run the jobs for 10 min to measure the ATT for a sufficient period. In addition, since the message size greatly affects the ATT, we experiment by increasing it from 64 KB to twice as much as 1,024 KB (1 MB).

The second comparative measure, a load balancing (*LB*), means how fairly the receiver tasks work. We compute the LB of each stream partitioning as Eq. (1).

$$LB(\%) = \left(1 - \frac{\sum_{k=1}^{n} \left| \begin{array}{c} Average\,no.\,of\,tuples - \\ Actual\,no.\,of\,messages\,emitted\,to\,receiver\,task\,k \end{array} \right|}{No.\,of\,total\,tuples}\right) \times 100 \qquad (1)$$

Equation (1) means the percentage of the difference between the number of messages when they are equally distributed to each receiver task and the number of actually received messages. Therefore, the higher the LB, the better the distribution, and the lower the LB, the worse the distribution, meaning that certain tasks more work. Since LB is important for the number of messages, not the message size, we here fix the size of the message small.

3.2 Application Configuration Cases

We design three job configuration cases to compare the performance of stateless stream partitioning in various distributed stream applications: (1) sender and receiver imbalance, (2) sender imbalance, and (3) receiver imbalance. We assume that receivers are deployed on all nodes so that stream partitioning takes good account of the task locality in all configurations. Table 1 shows the actual configurations to be used in the experiment in Sect. 4. These configurations make the most of our computing resources.

Case 1. Sender and receiver balance (*SRB*). The number of senders and receivers is balanced with the number of nodes. Here, the balance means that the same number of senders or receivers operate on all nodes, because the number of senders and receivers is a multiple of the number of nodes.

Case 2. Sender imbalance (*SI*). The number of senders is less than the number of nodes or is not a multiple of the number of nodes. Here, the number of senders deployed at each node is different, so that the number of messages sent by a particular node may be greater than the number of messages sent by other nodes.

Case 3. Receiver imbalance (*RI*). The number of receivers is not a multiple of the number of nodes. Here, a particular receiver may need to do more work than other receivers, depending on the stream partitioning.

4 Experimental Evaluation

In this section, we compare three stateless stream partitioning methods, SG, LoSG, and LAG on different network bandwidths. The hardware specifications are as follows: one Intel Xeon 2.4 GHz 8 Core and 16 GB RAM server for master node, and eight Intel Xeon 2.4 GHz 6 Core and 16 GB RAM servers for worker nodes. All nine servers are deployed in a distributed environment using EFM ipTIME T16000 network equipment for Ethernet, and Mellanox SwitchX®-2 MSX6012F-1BFS managed FDR 56 Gbps network equipment for InfiniBand. We configure four network bandwidths from two networking devices: 100 Mbps and 1 Gbps Ethernet and 40 Gbps and 56 Gbps InfiniBand.

Table 1 Three job configuration cases for the experiment

Case	# of worker nodes	# of sender tasks	# of receiver tasks
SRB	8	8	8
SI	8	4	8
RI	8	8	12

4.1 Comparison of Average Transmission Time

The first experiment compares the ATT of stream partitioning according to the network bandwidth in the three job configuration cases designed in Sect. 3.2. Figures 2, 3 and 4 show the ATT measured by varying the message size. First, LoSG always shows low ATT regardless of network bandwidth and message size. This is because LoSG communicates only within a node, so no network communication occurs. Conversely, the ATTs of SG and LAG increase as the message size increases, because both groupings require network communication between nodes. Especially, in Ethernet, the increase of LAG considering the locality is not greater than that of SG which does not consider the locality. In addition, SG has a surge in ATT, which makes it difficult to measure performance at 100 Mbps network bandwidth (Figs. 2a and 4a). This is because the ATT is much higher than the message creation period of 100 ms, and the job cannot operate normally. However, in InfiniBand, the ATT of SG is similar to the that of LAG, because the bandwidth of InfiniBand is so high that it quickly handles frequent communications of SG. In summary, when using high-bandwidth equipment, we can say that the performance loss is small even if simple SG is used instead of complex LAG.

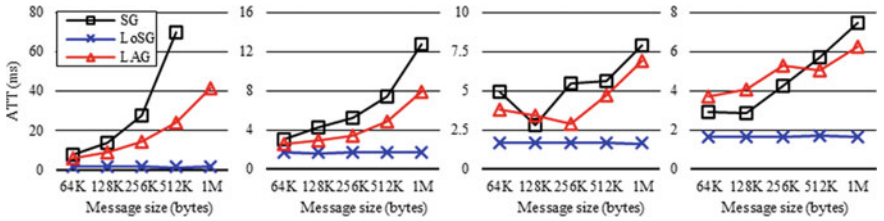


Fig. 2 Average transmission time in the sender and receiver balance

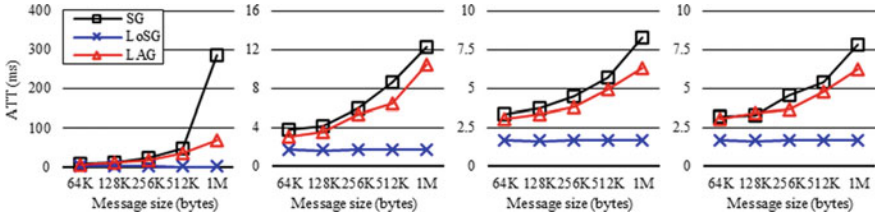
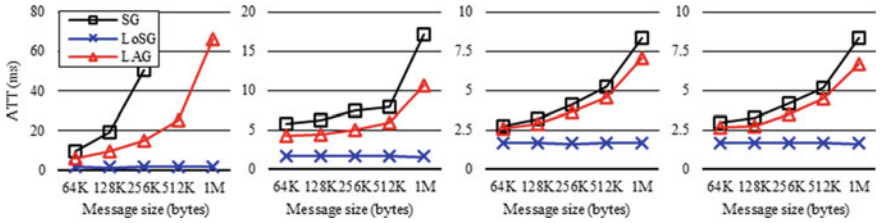


Fig. 3 Average transmission time in the sender imbalance



(a) 100 Mbps. (b) 1 Gbps. (c) 40 Gbps. (d) 56 Gbps.

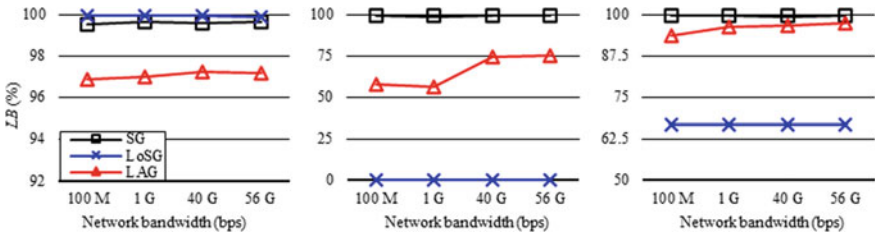
Fig. 4 Average transmission time in the receiver imbalance

4.2 Comparison of Load Balancing

The second experiment compares the LB of stream partitioning according to the network bandwidth in the three job configuration cases. Figure 5 shows the LB measured for each network bandwidth in three cases. First, in the SRB of Fig. 5a, the LB of SG is about 99.5%, that of LoSG is 99.9%, and that of LAG is close to 97.0%. SG and LoSG show very high LBs because the SG randomly sends messages to the receivers, and LoSG generates messages at the same rate. The LB of LAG is somewhat lower because there is a difference in network speed between nodes. However, about 97.0% is also very high, so load balancing has been done properly.

Second, Fig. 5b shows the LB measured in the SI. SG sends the same amount of messages to all receivers, so the LB is close to 100.0% regardless of network bandwidth. However, in LoSG, only four receiver tasks running on the same node as the sender task receive the message, and the remaining four receiver tasks do not receive the message. Thus, all experiments showed 0% of the LB. On the other hand, LAG, which alleviates the imbalance problem of LoSG, shows the LB of about 56.0–75.0%. In particular, the faster the network speed, the higher the LB. This means that a receiver task on a node that does not have a sender task will also process some messages.

Third, Fig. 5c shows the LB measured in the RI. SG shows an LB to 100.0% like the SI. Next, LoSG shows an LB of 66.7% because the receiver task of a node



(a) SRB. (b) SI. (c) RI.

Fig. 5 Load balancing measure of the receiver imbalance

that only operates on one receiver task performs twice as many tasks as another receiver task to be. Lastly, LAG shows a relatively high LB of about 94.0–97.0%, and the faster the network speed, the higher the result. As a result, LAG shows better performance in terms of load balancing as the network bandwidth increases, and better performance in terms of transmission bandwidth as the network bandwidth decreases.

5 Conclusions

In this paper, we benchmarked the stateless stream partitioning methods through extensive experiments on various network bandwidths. The experimental results could be summarized as follows:

(1) SG has the highest load balancing measure at all bandwidths and all balance cases. However, at low bandwidth, the larger the message size, the much higher the transmission time. Therefore, if the network bandwidth is low, SG is not recommended.
(2) LoSG is best for the sender and receiver balance. However, it is not recommended in imbalance cases, because its load balancing measure is very low.
(3) LAG compromises on the tradeoff relationship between SG and LoSG. That is, LAG has a lower transmission time than SG at low bandwidth, and a higher load balancing measure than LoSG in imbalance cases. Therefore, if we do not want to consider network bandwidths and balance cases, LAG is most recommended.

We believe that, using the experimental results of this paper, it is advantageous to select the correct stateless stream partitioning method according to network bandwidths, message sizes, and job configurations.

Acknowledgements This research was partly supported by Korea Electric Power Corporation (Grant number:R18XA05). This research was also partly supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2019R1A2C1085311).

References

1. Toshniwal A et al (2014) Storm @Twitter. In: Proceedings of the international conference on management of data, ACM SIGMOD, Snowbird, Utah, pp 147–156
2. Iqbal MH, Soomro TR (2015) Big data analysis: apache storm perspective. Proc Int J Comput Trends Technol (IJCTT) 19(1):9–14
3. Son S et al (2018) Locality aware traffic distribution in Apache storm for energy analytics platform. In: Proceedings of IEEE international conference on Big Data and smart computing, Shanghai, China, pp 721–724

4. Lu X et al (2013) High-performance design of hadoop RPC with RDMA over InfiniBand. In: Proceedings of the IEEE 42nd international conerence on parallel processing (ICPP), Lyon, France, pp 641–650
5. Huang J et al (2012) High-performance design of HBase with RDMA over InfiniBand. In: Proceedings of the IEEE 26th international parallel and distributed processing symposium (IPDPS), Shanghai, China, pp 774–785
6. EsperTech Esper, https://www.espertech.com/esper/.
7. Shoro AG, Soomro TR (2015) Big Data analysis: apache spark perspective. Proc Glob J Comput Sci Technol 15(1)
8. Chintapalli S et al (2016) Benchmarking streaming computation engines: storm, flink and spark streaming. In: Proceedings of the IEEE international parallel and distributed processing symposium workshops (IPDPSW), Chicago, IL, pp 1789–1792
9. van der Veen JS et al (2015) Dynamically scaling apache storm for the analysis of streaming data. In: Proceedings of the IEEE 1st international conference on Big Data computing service and applications, Redwood City, CA, pp 154–161
10. Batyuk A, Voityshyn V (2016) Apache storm based on topology for real-time processing of streaming data from social networks. In: Proceedings of the IEEE 1st international conference on data stream mining & processing (DSMP), Lviv, Ukraine, pp 345–349

Development of a Non-contact Autostereoscopic 3D Button Using Artificial Intelligence

Sang-Hee You, Min Hwang, Ki-Hoon Kim, and Chang-Suk Cho

Abstract This study presents the results of the development of a contactless button device that represents the button as an autostereoscopic vision and visually captures the position of the fingertip that presses the virtual button. To this end, a 3D stereoscopic expression module, a pointing location display module, an FPGA design for driving the modules, and an artificial neural network algorithm were developed. The pointing accuracy of the developed button device showed 99% accuracy in the recognition test in the laboratory.

Keywords Stereoscopic · Contactless · Virtual button · 3D · AI

1 Introduction

Pushbuttons, which have been used on most machines, have been around for a long time since industrialization began, but as modern three-dimensional technology rapidly evolves, it is time to replace them with new non-mechanical technologies. As a new button technology, the contactless button as the most promising candidate to replace the existing mechanical button is required in terms of security and hygiene. With the development of modern biometric technology, the need for security of touch-type buttons has emerged with the development of fingerprint recognition technology. In other words, the touch-type button has a problem that the password is easily exposed because there remains a trace of the contacts on the button, therefore there occurs a need to spread the non-contact type. In terms of hygiene, there is a desperate need for contactless buttons to prevent infection from highly infectious viruses or bacteria, such as the MERS virus in 2015 and the Corona virus that brought the 2020 pandemic. After experiencing the MERS crisis, which caused

S.-H. You · M. Hwang · K.-H. Kim · C.-S. Cho (✉)
Ivsys Co. Ltd., Osan-Si, Gyeonggi-Do 18101, Korea
e-mail: cscho@hs.ac.kr

C.-S. Cho
Division of Information and Telecommunication, Han-Shin University, Osan-Si, Gyeonggi-Do 18101, Korea

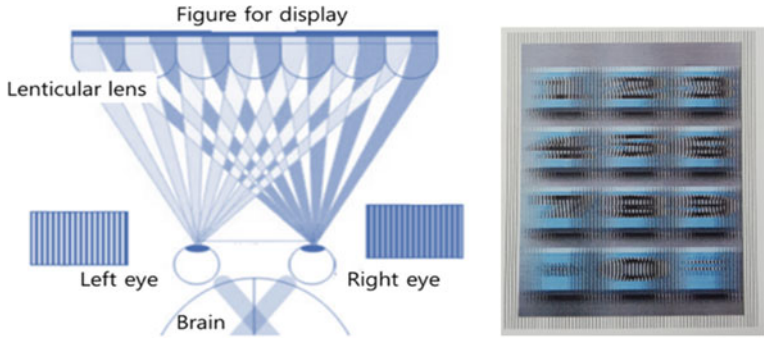
J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_10

Korean society to be shocked by the virus epidemic, we started to develop a contactless button. In order to design such a non-contact button, it was necessary to consider three points in terms of real demand. One was that the non-contact type should not have much difference compared to the mechanical type in terms of manufacturing cost, and the other was that the durability should be better than the mechanical type, and the last was that the consumer should be able to feel the point of pressing the button non-contactly. In consideration of the above points, the expression of a virtual button by autostereoscopic vision was envisioned in order to provide a feeling that the push button is on the button plate even in a contactless manner. To reduce costs and ensure stability, light-emitting and receiving sensors composed of infrared diodes were used, and the data obtained from the sensors are identified through a simple deep learning network.

A non-contact button using humidity sensor was reported [1] but its application field is more restricted than our vision button due to using humidity information. Various types of autostereoscopic vision technology have been developed since the 1970s, and their implementation has also diversified from cinema screens to LCD [2–4]. The most popular method is the implementation of autostereoscopic vision by making a multi-view using a lenticular lens, which is simple to implement and inexpensive. The lenticular method [5] was also used in this study. The recognition of the pointing position was performed using a simple deep learning network [6–8]. The structure and concept of deep learning was introduced in earnest from the 1980s, and is based on the construction of the middle layer extending from the initial artificial neural network. In the recognition experiment with this experimental device, there was no significant difference in recognition accuracy between a simplified deep learning network and a more complex deep learning network including CNN filters. Based on this, a deep learning network with a simple structure was used to recognize the pointing position. The recognition success rate of the non-contact virtual button was recorded over 99%, so that commercial success can be expected.

2 Button Device with Lenticular Display

The lenticular method distributes left and right images using the refraction of a lens which has semi-cylindrical columns arranged vertically on a screen. Figure 1 shows the principle of autostereoscopic vision using lenticular method. That is, when the convex lens-shaped semi-cylindrical lenses are arranged on a transparent plate vertically and the left and right images are alternately arranged under the plate, the images are distributed to the left and right eyes as shown in Fig. 1. Since the distributed left and right images are combined in the brain as a 3D form, the object is seen in stereoscopic vision. The lenticular method is the most widely used method in the field of autostereoscopic vision, and its implementation cost is also the cheapest, whereas if the target screen is slightly out of the main viewing angle, stereoscopic vision is not well recognized. The depth information of the virtual 3D by the lenticular lens is proportional to the thickness of the lenticular lens plate. The thickness of the lens plate



(a) Principles of Lenticular vision

(b) Virtual 3D button set with lenticular lens

Fig. 1 Lenticular autostereoscopic vision and the button set in lenticular display

used in this experiment is 4 mm, so the maximum depth of the virtual 3D is 1.8 cm. On the other hand a touch panel key system with lenticular lens was reported as an autostereoscopic button in Fujitsu (Patent application number: 10–2019-0,038,210, Korea). When a button in the system is expressed, then the touch panel detects the input as touch panel sensor. This method is a heavy and expensive system due to using touch screen, and the problem of recognizing the pointing position of a finger as non-contact has not been solved. Comparing to our proposed method the autostereoscopic touch panel has not been put into practical use yet. Comparing to those methods our proposed button does not require LCD screen or special sensor system, so that it has the merits of very low implementation cost, lightweight and contactless style.

The autostereoscopic button device is composed of a display part with lenticular lens, an infrared sensor part, a backlight part for each button, and a driving unit including artificial intelligence. Figure 2 shows the configuration of the button device. The button picture reconstructed by dividing a button photo into multi-viewpoints

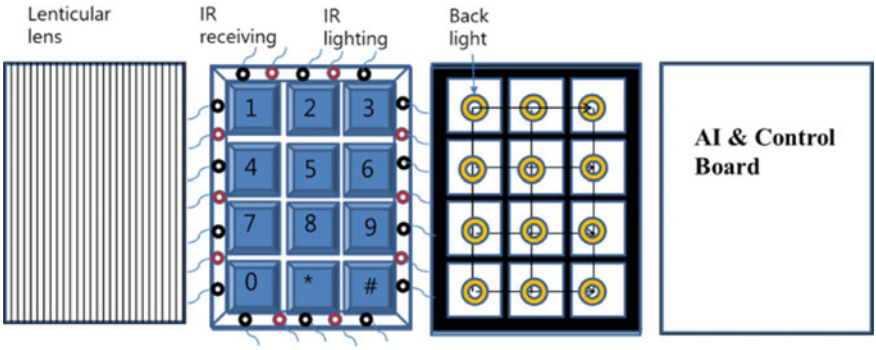


Fig. 2 Non-contact 3D virtual button device

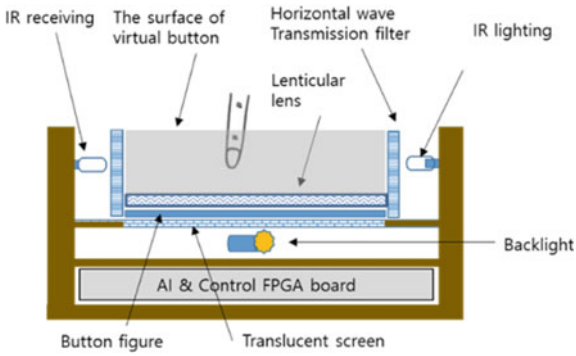


Fig. 3 The button device cross-section

is mounted under the lenticular lens, and a backlight is disposed on the back side of the picture. The backlight serves to indicate the position where a virtual button is pointing. Without such a backlight a user can't recognize whether a button has been pressed or not because it is a non-contact method with virtual button. At the each edge of the horizontal and vertical axis of the button device, infrared diodes are arranged.

Figure 3 shows the cross-section and pointing situation of the button device. When a fingertip enters into the virtual button floated into the air by the autostereoscopic vision, the AI system recognizes the button as pressed. Artificial intelligence is used to determine whether a finger has entered into the virtual button image. Figure 3 shows a cross-section of the button. The infrared diodes consist of light emitting sensors and light receiving sensors, and are arranged in three pairs on the upper and lower sides and four pairs on the left and right sides. In the case of the infrared light emitting diode, a total of 14 are arranged on the top, bottom, left, and right of the button device, and light is sequentially turned on when emitting. The infrared light emitting diode here emits light including wavelengths from near infrared to far infrared. The light-receiving diode does not selectively select a specific wavelength, but detects it integrally, including infrared wavelengths from near infrared to far infrared. The infrared light emitting sensors arranged for each position of the button device sequentially emit light clockwise one by one, so that the light receiving sensors detect light emission simultaneously from every light emission. Even if the light is emitted sequentially, it is instantaneous when observed, so it appears as a simultaneous light on the photo.

3 Pointing Algorithm Using Deep Learning

The deep learning network learns the 14 × 14 pixel map of images obtained by the receiving sensors shown in Fig. 4. The control board sequentially lights up 14 light emitting diodes and obtains a 14 × 14 pixel gray image composed of 14 gains from light-receiving diodes. The obtained gray image is learned in a deep learning

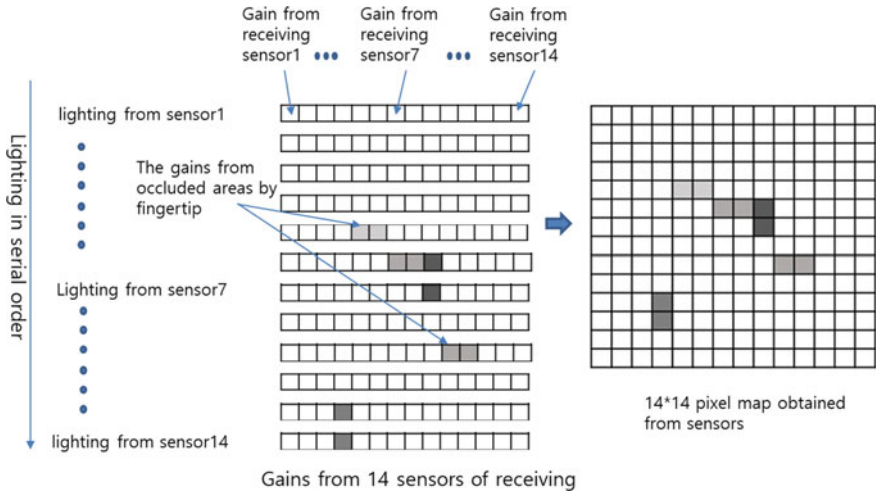


Fig. 4 Gains from 14 receiving diodes and image map obtained

network on the control board and outputs the pressed button information. That is, when one light-emitting diode is turned on, 14 light-receiving diodes around the button set simultaneously receive infrared light to output the detected value. The light-receiving sensor in the area blocked by a fingertip outputs a small value and the sensor unblocked does a large value. By this principle, it is possible to capture the position of the fingertip. In Fig. 4 it can be seen a 14 × 14 image composed of light emission and light reception for each sensor. Figure 5 shows the lighting control and

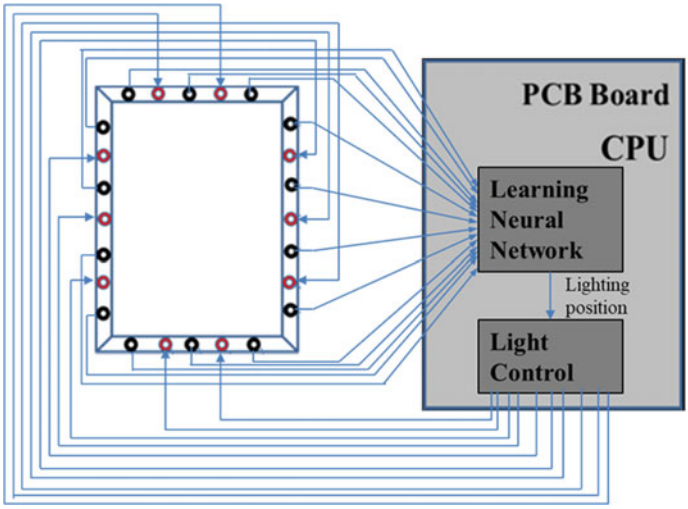


Fig. 5 The lighting control and learning system on PCB board

learning system on PCB board. The light-receiving diodes blocked by the fingertips are relatively clearly distinguished. Thus it can be determine the fingertip position by establishing a linear equation, so that the button pressed can be obtained without a deep learning network. However the problem is that the use of a linear relational expression is very vulnerable to changes in the lighting environment, so that we decided to use a deep learning network in spite of its simple input pattern.

As described above in Sect. 3, the gray image for input is an image created by 14 light emitting diodes sequentially emitting and 14 light receiving diodes recording the input value. Therefore, a relatively simple network with two middle layers was used because the input image has relatively simple and clear features.

Figure 6 shows the neural network used in this study. Since the deep learning network in Fig. 7 receives a single 14 × 14 pixel image as an input, the number of input nodes is 196, and the middle layers are composed of 100 and 50 nodes respectively, and the output layer is composed of 12 nodes. This network includes

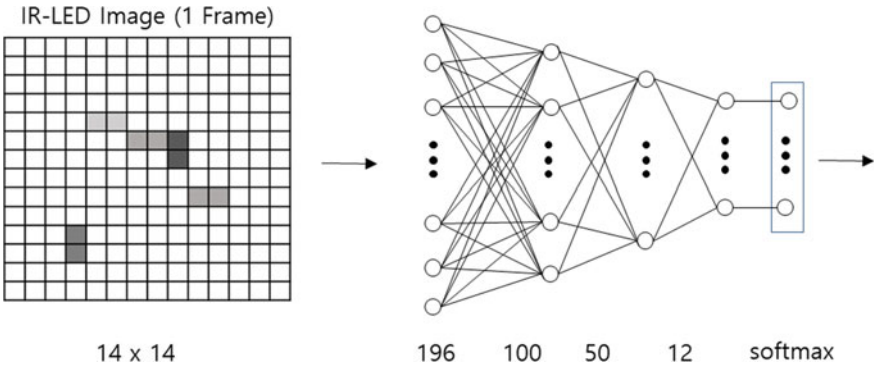
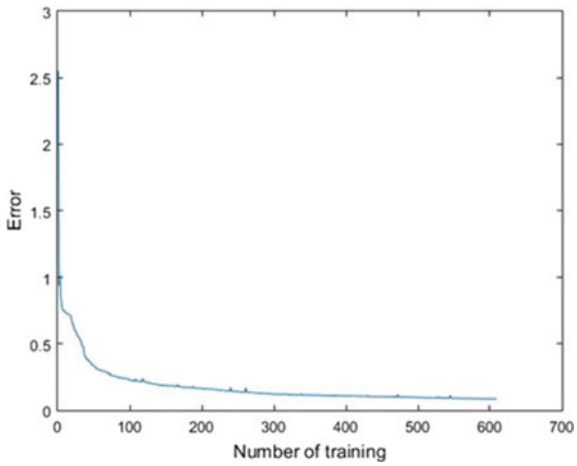


Fig. 6 Simple deep learning network for the fingertip-pointing

Fig. 7 Error according to learning iteration



no convolution filers in layers due to its simple inputs. The two hidden layers used sigmoid function as an activation function but did not use the activation function in output layer (i.e. Bypass). The value of the output node was converted to a stochastic value (between 0 and 1) by using SOFTMAX in the cross-entropy method. The SOFTMAX function is an activation function generally used in the output layer of a deep learning model for classification of 3-class or higher. The SOFTMAX function expands the deviation of each value when there are K values making the larger value relatively larger and the smaller value relatively smaller, and then normalizing it. The final output layer here is 12 because the number of buttons consists of 0–9 and 2 special characters. As a cost function, MSE (Mean Square Error) between the output value and the correct answer was used.

4 Experimental Result

For the experiment, a deep learning network was trained with 100 learning images extracted per position by placing a finger at the position of 12 key areas of the 3D virtual button. As a learning environment, a teaching button set for inputting manually target value was designed for easy learning and connected to the stereoscopic button device during learning. When the network is being learned, a position pointed on the stereoscopic button is identified to the target position inputted on the teaching button set. If the button set is learned for 600 epochs which have 1200 images respectively under an environment, it shows 99% or more correct answers in our experiments. Figure 8 shows the error (MSE: Mean Square Error) according to 600 iterations. The MSE was computed in Eq. 1.

$$\mathrm{MSE} = \frac{\sum \delta(i)^2}{n}, \begin{cases} \text{if } i = T \text{ then } \delta(i) = 0 \\ \text{else } i \neq T \text{ then } \delta(i) = 1 \end{cases}, i : \text{Inputted value}, T : \text{Target} \quad (1)$$

Reliability is very important in the button, so we have to go through two verification processes as shown in Fig. 8. That is, if the results of the two recognitions match, the results are output, and if they do not match, the recognition process is

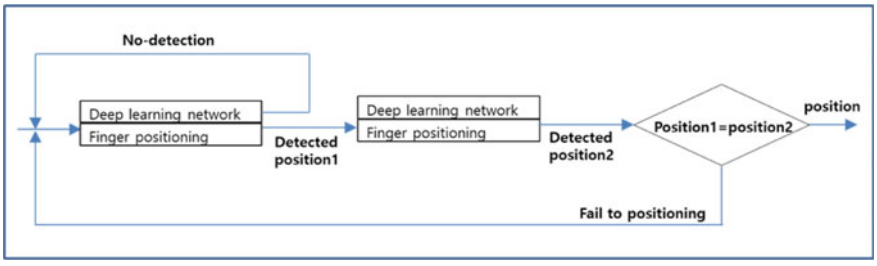


Fig. 8 Double verification method to improve reliability

performed again. There is no problem in using this repetitive process because the time difference is small enough that the user cannot feel it.

5 Conclusion and Future Work

This study presented the design of a stereoscopic virtual button as non-contact, the configuration of a button sensor system, and the design of a simple deep learning network to recognize it. If there is a clear design of the input data composition, it proved that the deep learning network needs not be so complicated. The production of contactless devices according to the recent social distance is expected to gradually expand its scope. As a result of this study, the proposed new concept button has a feature of high reliability of over 99% and simple hardware configuration with relatively low cost. In the future, this device will be released after additional verification processes for commercialization.

Acknowledgements This work was partly supported by the Hansin University Grant and Technology development Program of MSS [S2758698] and of MSS [S2721154].

References

1. Mondal S, Kim SJ, Choi C-G (2020) Honeycomb-like MoS2 nanotube array-based wearable sensors for noninvasive detection of human skin moisture. ACS Appl Mater Interfaces 17029–17038
2. Dodgson NA (2005) Autostereoscopic 3D displays. IEEE Comput 38(8):31–35
3. Tyler CW, Clarke MB (1990) Autostereogram. In: Proceedings SPIE 1256, stereoscopic displays and applications
4. Ezra D, Woodgate GJ, Omar BA, Holliman NS, Harrold J, Shapiro LS (1995) New autostereoscopic display system. In: Proceedings SPIE 2409, stereoscopic displays and virtual reality systems II
5. Bourke P (2014) Autostereoscopic lenticular images. Lenticular Prints. https://paulbourke.net/stereographics/lenticular/
6. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. Advances in neural information processing systems, vol 25. NIPS
7. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. In: Computer vision and pattern recognition, vol 6
8. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. IEEE CVPR, pp 770–778

Automatic Computing Device Selection Scheme Between CPU and GPU for Enhancing the Computation Efficiency

Geunmo Kim, Sungmin Kim, Jinsung Cho, Jeong-Dong Kim, and Bongjae Kim

Abstract Recently, cluster computer environments using multiple CPUs or multiple GPUs have been widely used to accelerate computation such as in the field of big data or artificial intelligence. In this paper, we propose an automatic computing device selection scheme between CPU and GPU by estimating the expected execution time based on trained deep learning model. Based on the experimental results, our scheme can select the computing device dynamically to reduce the execution time and maximize the computing efficiency.

Keywords CPU · GPU · Computing efficiency · Device selection · Computing acceleration

1 Introduction

Cluster computing environments using multiple CPUs and GPUs are frequently used to accelerate various computing operations or jobs [1, 2]. The CPU and GPU are structurally different. Due to structural differences between CPU and GPU, the processing efficiency in terms of the execution time depends on the amount of processing [3]. In the case of a GPU, to perform a computing operation, data required for the computation must exist in the global memory of the GPU. Therefore, to perform the computing

G. Kim · S. Kim · J. Cho · J.-D. Kim · B. Kim (✉)
Division of Computer Science and Engineering, Sun Moon University, 70, Sunmoon-ro 221beon-gil, Tangjeong-myeon, Asan 31460, Republic of Korea
e-mail: bjkim0422@gmail.com

G. Kim
e-mail: rootmo96@gmail.com

S. Kim
e-mail: kimsungmin.dev@gmail.com

J. Cho
e-mail: cjs97111@gmail.com

J.-D. Kim
e-mail: kjdvhu@gmail.com

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_11

operation on the GPU, data copy overhead occurs from the main memory to the global memory of the GPU. Besides, there is an overhead to bring the result of the computation from the global memory of the GPU to the main memory [4]. In short, if the computational load is small, the CPU may be more efficient than the GPU due to memory copy overheads. For the above reasons, there is a need for a technique that can predict the expected execution time of the CPU and GPU and automatically select the computing device according to the given computation workload.

In this paper, we proposed an automatic computing device selection scheme between CPU and GPU by estimating the expected execution time. In this paper, we focus on the pooling operation. The pooling operation is commonly used in a deep learning model [5]. It is also used in other data analysis models. In our scheme, we use a deep learning model to estimate the expected execution time. The pooling operations of the CPU version and the GPU version were implemented respectively. Actual execution time according to the size of the filter, the size of the matrix, the number of CPU cores used, and the number of GPU devices used are extracted. Our deep learning model which is used to estimating the execution time was trained using the actual extracted execution time, the size of the filter, the size of the matrix, the number of CPU cores used, and the number of GPU devices used. Based on the trained execution time prediction model, we estimate the expected execution times when using the CPU and GPU respectively. Then, our scheme minimizes the execution time of the given pooling operation by selecting the computing device with a small expected time between CPU and GPU. We implemented the proposed technique to confirm that the proposed technique is effective. Based on the experimental results, our scheme can select the computing device dynamically to minimize the execution time and maximize the computing efficiency.

The rest of this paper is as follows. In Sect. 2, we will explain our automatic computing device selection scheme in detail. We will show the experiment results of our scheme in Sect. 3. Finally, Sect. 4 concludes this paper with some future works.

2 Automatic Computing Device Selection Scheme

Table 1 shows the pseudocode of our automatic computing device selection scheme. *CPU_expected* means the expected execution time based on our deep learning model when using the CPU device to perform the given pooling operation. Similarly, *GPU_exepected* denotes the expected execution time when using the GPU device. The *filter_size*, *matrix_size*, *cpu_count*, and *gpu_count* denote the size of the filter, the size of the matrix, the number of CPU cores used, and the number of GPU devices used, respectively.

In our scheme, the deep learning model is used to predict the execution time. *CPU_expected* and *GPU_exepected* are predicted by each different deep learning model. In other words, there is a deep learning model for predicting *CPU_expected* and a deep learning model for predicting *GPU_exepected*. If the *CPU_expected* is greater than the *GPU_exepected*, our scheme calls the *do_pooling_with_GPU()*.

Table 1 The pseudocode of our automatic computing device selection scheme

```
do_pooling()
{
        CPU_expected = Calculate_CPU_Expected(filter_size, matrix_size, cpu_count)
        GPU_expected = Calculate_GPU_Expected(filter_size, matrix_size, gpu_count)

        if CPU_expected > GPU_expected then
                call do_pooling_with_GPU();
        else
                call do_pooling_with_CPU();
}
```

In the opposite case, our scheme calls the *do_pooling_with_CPU()*. In this way, our scheme can select a computing device between CPU and GPU to enhance computational efficiency.

Figure 1 shows the layer structure of our execution time prediction model. The Tensorflow 1.12.0 and the Keras 2.2.4 were used to implement our execution time prediction model. As shown in Fig. 1, the size of the filter, the size of the matrix, and the number of GPUs or CPUs used are input to output the predicted execution time. The model outputs the predicted execution time in seconds.

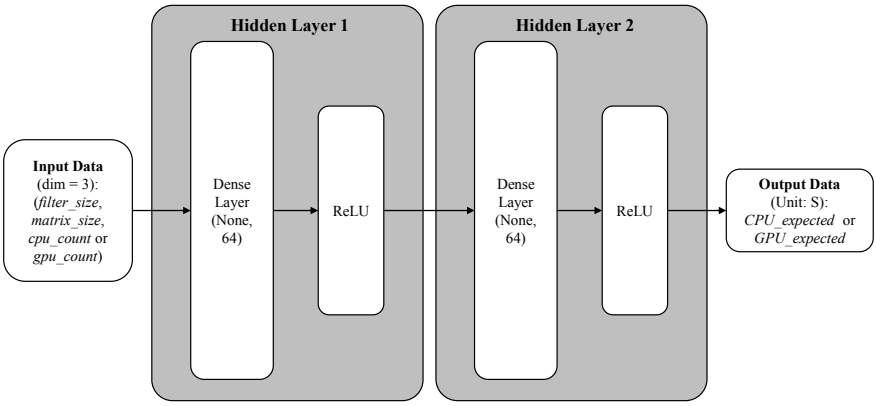


Fig. 1 The layer structure of our execution time prediction model

Table 2 Cluster computing environment used in the performance evaluation

Features	Computing Node 1	Computing Node2
CPU	Intel i7-8700 (6 cores, 3.2 GHz)	Intel i5-7500 (4 cores, 3.4 GHz)
RAM	16 GB	8 GB
GPU	NVIDA GTX 1050	
GPU Base Clocks	640 MHz	
GPU Memory Clocks	1354 MHz	
GPU Memory Capacity	2 GB GDDR5	
OS	Ubuntu 16.04.6 LTS	
CUDA Version	CUDA Toolkit 9.0	
MPI Version	MVAPICH 2–2.3.1	
CuDNN Version	Version 7.0.5	

3 Performance Evaluations

In this section, we will describe the experimental environment. The performance evaluation results of the proposed technique will be described and discussed.

3.1 Experimental Environment

Table 2 shows the experimental environment to evaluate the proposed automatic computing device selection scheme based on multiple CPUs and multiple GPUs with MPI. The MPI (Message Passing Interface) is a programming interface (library) used to exchange information or computing jobs between compute nodes and is widely used for cluster computing [6, 7]. In the experiment, we compared the performance of our scheme with two schemes; CPU only scheme and GPU only scheme. In the case of the CPU only scheme, only CPU is used for pooling operation. Similarly, only GPU is used for pooling operation in case of the GPU only scheme. As described above, the proposed scheme can dynamically select the CPU and GPU using the execution time prediction model.

3.2 Experimental Results

Figure 2 shows the execution time results of CPU only, GPU only, and our scheme when the size of the filter is 2×2, the number of CPU cores used is 4, and the number of GPU devices used is 2). As shown in Fig. 2, the proposed scheme has a

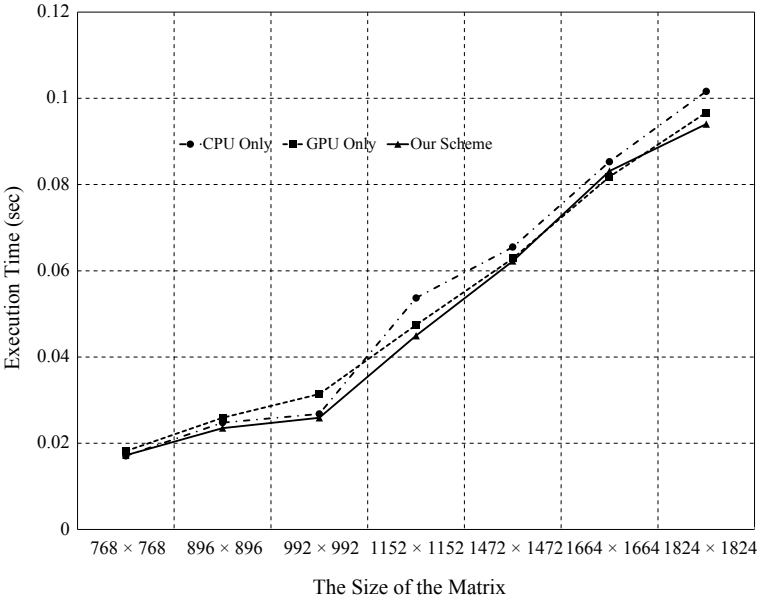


Fig. 2 Execution time results of CPU only scheme, GPU only scheme, and our scheme (the size of the filter = 2 × 2, the number of CPU cores used = 4, and the number of GPU devices used = 2)

shorter execution time compared to the CPU-only or GPU-only scheme. The average execution time of CPU-only scheme is about 0.054 s. The average execution time of GPU-only scheme is about 0.052 s. The average execution time of the proposed scheme is 0.05 s. If we compare our scheme with the CPU-only scheme, the performance in terms of the execution time is improved by about 8%. The degree of the performance improvement increases as the load of a given pooling operations increases. Based on the experimental results, we can confirm that our scheme can reduce the execution time and maximize the computing efficiency.

4 Conclusions and Future Works

In this paper, we proposed a dynamic computing device selection scheme between CPU and GPU by considering a given pooling job. According to the experimental results, the proposed scheme can minimize the execution time by dynamically selecting the computing device between CPU and GPU according to the load of the given pooling operation. Based on the experimental results, the performance of our scheme in terms of the execution time is improved by about 8% when compared to the CPU only scheme.

In future works, we will improve the execution time prediction model to support more diverse computations such as convolution operation. We will also improve the

proposed technique to support not only CPU or GPU devices but also devices that can support neuromorphic computing.

Acknowledgements This work was supported by Institute for information and communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No. 2019-0-00708, Integrated Development Environment for Autonomic IoT Applications based on Neuromorphic Architecture).

References

1. Chen C, Li K, Ouyang A, Li K (2018) FlinkCL: an openCL-based in-memory computing architecture on heterogeneous CPU-GPU clusters for Big Data. IEEE Trans Comput 67(12):1765–1779
2. Chen C, Li K, Ouyang A, Zeng Z, Li K (2018) Gflink: An In-memory computing architecture on heterogeneous CPU-GPU clusters for Big Data. IEEE Trans Parallel Distrib Syst 29(6):1275–1288
3. Syberfeldt A, Ekblom T (2017) A comparative evaluation of the GPU vs the CPU for parallelization of evolutionary algorithms through multiple independent runs. Int J Comput Sci Inf Technol (IJCSIT) 9(3):1–14
4. Kim B, Jung J, Min H, Heo J, Jung H (2017) Performance evaluations of multiple GPUs based on MPI environments. In: Proceedings of the international conference on research in adaptive and convergent systems. ACM, pp 303–304
5. Lee CY, Gallagher PW, Tu Z (2016) Generalizing pooling functions in convolutional neural networks: mixed, gated, and tree. In: Artificial intelligence and statistics, pp 464–472
6. Lu J, Zhang K, Chen M, Ma K (2013) Implementation of parallel convolution based on MPI. In: Proceedings of 2013 3rd international conference on computer science and network technology. IEEE, pp 28–31
7. Li X, Zhang G, Huang HH, Wang Z, Zheng W (2016) Performance analysis of Gpu-based convolutional neural networks. In: 2016 45th international conference on parallel processing (ICPP). IEEE, pp 67–76

Experimentation of Human Activity Recognition by Using Accelerometer Data Based on LSTM

Phataratah Sa-nguannarm, Ermal Elbasani, Bongjae Kim, Eung-Hee Kim, and Jeong-Dong Kim

Abstract The majority of studies on Human Activity Recognition (HAR) were conducted utilizing video and/or image data which caused critical problems such as privacy exposure and storage capacity issues. To tackle these issues, we propose a Long Short-Term Memory (LSTM) architecture for HAR taking accelerometer data acquired from smartphones as its input values. The experimental results showed the outstanding performance of the proposed model against a convolutional neural network (CNN) based model which is one of the state-of-art methods for HAR. More precisely, it was observed that the accuracy of the LSTM based model is about 12% superior to that of the CNN-based model.

Keyword Human activity recognition · Long short-term memory · Convolutional neural networks

1 Introduction

The life monitoring or lifelogging model aims to improve people's quality of life and wellness care [1, 2]. Life monitoring was applied in many fields such as the

P. Sa-nguannarm · E. Elbasani · B. Kim · J.-D. Kim (✉)
Division of Computer Science and Engineering, Sun Moon University, Asan, South Korea
e-mail: kjdvhu@gmail.com

P. Sa-nguannarm
e-mail: phataratah.sanguan@gmail.com

E. Elbasani
e-mail: ermal.elbasani@gmail.com

B. Kim
e-mail: bjkim@sunmoon.ac.kr

E.-H. Kim
Department of Global Software Engineering, Sun Moon University, Asan, South Korea
e-mail: ehkim@sunmoon.ac.kr

J.-D. Kim
Genome-based BioIT Convergence Institute, Sun Moon University, Asan, South Korea

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_12

medical field, transportation detection, and digital memory [2–4]. However, existing models used cameras for recording daily lifelog but pictures have some problems. They need large capacity and expensive storage [5]. And one problem of cameras based life monitoring is the user's privacy issue [6]. Some researchers established the model that uses wearable sensors such as accelerometers, gyroscope, but it is still inconvenient in a real-life environment [6, 7]. This work presents life monitoring by using time series data from smartphone sensors such as accelerometers. This has led to more individual data analysis perspective and low cost.

Researchers have tried many approaches for aiming for the best performance in Human Activity Recognition (HAR) tasks for lifelogging monitoring such as machine learning techniques with hand-crafting features, and deep learning techniques by using data from various sensors [8–10]. To process and analyzed these data, recently deep learning methods Recurrent Neural networks (RNN) with also the utilization of Long Short-Term Memory (LSTM) and as well the Convolutional Neural Networks (CNN), have provided state-of-art solutions for the recognition of activities in several natures of studies. The purpose of this study is to compare the performance of the LSTM and CNN model by using time series data from accelerometers embedded in a smartphone to personalized, fully automated, and low-cost implementation in HAR lifelogging monitoring.

This paper is organized as follows. In Sect. 2 is described the related work. In Sect. 3 experiments, we compared the accuracy of the LSTM model with the accuracy of the CNN model. Finally, concluding with Sect. 4.

2 Related Work

Researchers established lifelogging monitoring by using data from a wearable accelerometer and gyroscope [6, 7, 9]. Lee et al. [6] established a daily lifelogging system using an accelerometer because it is small, wearable, and saves energy. They used an accelerometer with a laptop computer to gather the data. Park et al. [7] implemented the lifelog monitoring model by using their proposed wearable sensor. They constructed the hardware for gathering the lifelog data. The hardware has a gyroscope, heart rate sensor, Bluetooth, and tact switch. Besides, get the GPS data from the mobile device. That seems a laboratory framework and uncomfortable to use in real life. Kwapisz et al. [9] firstly tried to acquire accelerometer time-series data from Android smartphones for HAR tasks because the Android operating system is free, uncomplicated to program, and become pervasive in the cell phone market. This research became benchmark work and a dataset for HAR tasks until now.

In past years, researchers have used machine learning methods to perform HAR tasks [9]. The works that used machine learning methods were accurate but it took

time in a feature extraction step. In the deep learning era, researchers analyzed deep learning for improving the performance of HAR tasks. Chen et al. [11] presented the LSTM Networks to perform the HAR task but needed statistical feature extraction. Zhang et al. [12] presented the deep neural network with a regular vine copula approach to classifying data gathered from smartphones. They are also required to extract high-level features from the sensors data. However, the previous works presented deep learning approaches needed to extract features from sensors data. Hence, this paper experiments the LSTM model without the feature extraction step for the HAR task and compares it with the CNN method.

3 Experiments

This study proposes a HAR that compares the performance of the LSTM with the CNN model (that was published on the website [13, 14] by order) and is considered to be an extension of our previous work [7] for leveraging in the same domain.

The LSTM and CNN models were established by Python version 3.6.9, TensorFlow version 2.0, and Keras version 2.3.1. Softwares are installed in Linux Ubuntu operating system (18.04).

3.1 Data Set Description

Two benchmark datasets that use smartphones are Wireless Sensor Data Mining (WISDM) [9] and Heterogeneity Activity Recognition Dataset (HHAR) [15]. The WISDM dataset has been released by WISDM Lab. The data was obtained from a smartphone in a plant's pocket which has an embedded 3-axis accelerometer with a sampling rate of 20 Hz (1 sample every 50 ms) of 36 subjects. 1 subject performed 6 activities like walking, jogging, sitting, standing, upstairs, downstairs. This dataset has 1,098,207 samples. The HHAR dataset obtained from an embedded 3-axis accelerometer sampled at the highest frequency in smartphone 8 models: 2 Samsung Galaxy S3 mini, 2 Samsung Galaxy S3, 2 LG Nexus 4, 2 Samsung Galaxy S+. Dataset has 9 subjects to perform 6 activities like biking, sitting, standing, walking, going upstairs, going downstairs. This dataset has 1,048,576 samples.

3.2 Experimental Setup

LSTM. The LSTM model [13] begins with a fully-connected layer with the Rectified Linear Unit (ReLU) activation function. Weight and bias for each neuron in this layer were randomized between the time step of input and number of features. Then, the output was calculated by 2 layers of LSTM with forget bias is 1.0 and went through a

recurrent layer. The last layer applied a softmax function with Adam optimizer. Also, applied l2 loss function in the last layer. Data preprocessing step starts with segments each axis with 20-step overlapping, 200-time steps. Then, split segmented data into a training set 70% and testing set 30%. The final shape of the train and test set are 3 features, 200-time steps. The number of samples in the WISDM dataset is 54,091 samples and the HHAR dataset is 46,895 samples. To estimate the performance of the model by changing hyperparameters in the model. Firstly, the varying number of layers in a fully-connected neural network from 1 to 3 layers, the number of neurons in each layer is 64 and 100 neurons, and the number of layers in the LSTM layer from 1 to 3 layers. Also, the learning rate from 0.0020 to 0.0030, L2 loss function value is 0.0015 and 0.0020, and the epoch from 10 to 300. Before training the model, split the training set by K-fold validation ($k = 5$) into the training set and the validating set.

CNN. The CNN model [14] processes the 3-axis accelerometer data starting in a 2-dimension convolutional layer with 256 convolutional filters. The filter size is 2×2 and applied ReLU to the data. The second layer is a 2-dimension max-pooling size 2×2 and a dropout ratio of 0.2. Flattening the output for reducing the dimension of feature maps to apply the dense layer number 1 and 2 that has 256 neurons each layer with the ReLU activation function. The last dense layer has several neurons according to the number of classes with softmax activation function and Adam optimizer. Data preprocessing step starts with segments at each axis with 50% overlapping, 90-time steps. Then split segmented data into a training set 70% and testing set 30%. The number of samples in the WISDM dataset is 15,048 samples and the HHAR dataset is 16,318 samples. Hyperparameter optimizations varied kernel size from 2×2 to 3×3, pooling window size from 1×1 to 2×2, the number of filters is 128 and 256, the number of neurons in fully-connected layers is 128 and 256. Included, varying batch sizes are 32, 64, 1024, and dropout rates are 0.3 and 0.4. Before training the model split the training set into the training set 70% and the validating set 30%.

3.3 Experimental Results

The highest accuracy of the LSTM model on the WISDM and the HHAR datasets are 95.82 and 95.17% by order. The highest accuracy of the CNN model on the WISDM and the HHAR datasets are 88.39 and 83.58% by order. The average accuracy of the LSTM model on the WISDM dataset is 95.57% and the HHAR dataset is 94.38%. The average accuracy of the CNN model on the WISDM dataset is 87.06% and the HHAR data set is 80.93%. That shows the average accuracy of the LSTM model on the WISDM and HHAR datasets that are higher than the CNN model 10.46 and 13.45% by order.

Also, evaluation metrics such as precision, recall, and F-1 score of each model with each dataset follow Gupta et al. [16] shown in Table 1 and 2 by order. The average precision, recall, and F1-score of the LSTM model on WISDM are higher than the CNN model 0.126, 0.112, and 0.119 by order. The average precision, recall,

Table 1 Performance of LSTM and CNN on the WISDM dataset

Activity	LSTM			CNN		
	Precision	Recall	F1-score	Precision	Recall	F1-score
Downstairs	0.939	0.943	0.941	0.697	0.629	0.662
Jogging	0.994	0.995	0.995	0.958	0.974	0.966
Sitting	0.992	0.985	0.989	0.963	0.981	0.972
Standing	0.984	0.996	0.990	0.869	0.973	0.918
Upstairs	0.942	0.934	0.938	0.663	0.695	0.679
Walking	0.991	0.991	0.991	0.940	0.921	0.930
Average	0.974	0.974	0.974	0.848	0.862	0.855

Table 2 Performance of LSTM and CNN on the HHAR dataset

Activity	LSTM			CNN		
	Precision	Recall	F1-score	Precision	Recall	F1-score
Bike	0.994	0.991	0.993	0.952	0.949	0.951
Sit	1.000	0.999	1.000	0.921	0.780	0.845
Stairsup	0.973	0.901	0.935	0.822	0.834	0.828
Stairsdown	0.917	0.974	0.945	0.843	0.900	0.871
Stand	0.998	0.999	0.999	0.943	0.872	0.906
Walk	0.970	0.978	0.974	0.926	0.860	0.787
Average	0.975	0.974	0.974	0.901	0.866	0.865

and F1 score of the LSTM model on the WISDM dataset are higher than the CNN model 0.074, 0.108, and 0.109 by order.

The previous method on WISDM is the LSTM Networks [11] reached 92.10% overall accuracy. The previous methods on the HHAR dataset are a regular vine copula-based approach [12] gained 88.3% precision. The LSTM model on the WISDM and HHAR datasets achieved 3.72% overall accuracy and 9.2% precision higher than previous works.

4 Conclusion

This paper presents experimental results of HAR approaches for daily life monitoring architecture. The performance of the LSTM model compared with the CNN model using time-series data from accelerometers in smartphones. The LSTM model has accuracy on benchmark datasets over 90% which is better than the CNN model 11.96% overall accuracy. Accuracy of the LSTM model is better than previous works that presented deep learning for HAR task 3.72% overall accuracy on the WISDM

dataset and 9.2% overall precision on the HHAR dataset. Hence, the LSTM model is more satisfactory than the CNN model with time-series data from accelerometers in smartphones.

Acknowledgements This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT)(No. 2019R1F1A1058394).

References

1. Biswas J, Tolstikov A, Jayachandran M, Foo SF, Victor PW, Aung A, Phua C, Huang W, Shue L, Gopalakrishnan K, Lee J-E, Yap P (2010) Health and wellness monitoring through wearable and ambient sensors: exemplars from home-based care of elderly with mild dementia. Ann Telecommun Annales des Télécommun 65:505–521
2. Yan H, Huo H, Xu Y, Gidlund M (2010) Wireless sensor network based e-health system—implementation and experimental results. IEEE Trans Consum Electron 56:2288–2295
3. Hemminki S, Nurmi P, Tarkoma S (2013) Accelerometer-based transportation mode detection on smartphones. In: Proceedings of the 11th ACM conference on embedded networked sensor systems, pp 1–-14. Association for Computing Machinery, Roma, Italy
4. Harvey M, Langheinrich M, Ward G (2016) Remembering through lifelogging: a survey of human memory augmentation. Pervasive Mob Comput
5. Wang P, Smeaton AF (2013) Using visual lifelogs to automatically characterize everyday activities. Inf Sci 230:147–161
6. Lee M-W, Khan AM, Kim T-S (2011) A single tri-axial accelerometer-based real-time personal life log system capable of human activity recognition and exercise information generation. Pers Ubiquit Comput 15:887–898
7. Park MG, Hwang JH, Putri Listio SW, Kim JD (2019) Design and implementation of Lifelog monitoring model using wearable sensors. In: 2019 international conference on digital contents (DigiCon-19)
8. Bayat A, Pomplun M, Tran DA (2014) A study on human activity recognition using accelerometer data from smartphones. Proc Comput Sci 34:450–457
9. Kwapisz JR, Weiss GM, Moore SA (2011) Activity recognition using cell phone accelerometers. ACM SIGKDD Explor Newsl 12:74–74
10. Yao S, Hu S, Zhao Y, Zhang A, Abdelzaher T (2017) DeepSense: a unified deep learning framework for time-series mobile sensing data processing. In: Proceedings of the 26th international conference on world wide web, pp 351–360. International world wide web conferences steering committee, Perth, Australia
11. Chen Y, Zhong K, Zhang J, Sun Q, Zhao X (2016) LSTM networks for mobile human activity recognition. In: Presented at the 2016 international conference on artificial intelligence: technologies and applications
12. Zhang S, Geng B, Varshney PK, Rangaswamy M (2019) Fusion of deep neural networks for activity recognition: a regular vine copula based approach. arXiv:1905.02703 [EESS]
13. Venelin V (2020) Human activity recognition using LSTMs on Android—TensorFlow for Hackers (Part VI). Last accessed 20 Feb 2020
14. Human Activity Recognition Using Convolutional Neural Network in Keras, https://github.com/Shahnawax/HAR-CNN-Keras. Last accessed 20 Feb 2020
15. Stisen A, Blunck H, Bhattacharya S, Prentow TS, Kjærgaard MB, Dey A, Sonne T, Jensen MM (2015) Smart devices are different: assessing and MitigatingMobile sensing heterogeneities for activity recognition. In: Proceedings of the 13th ACM conference on embedded networked sensor systems, pp 127–140. Association for Computing Machinery, Seoul, South Korea

16. Gupta R, Gurrin C (2019) Activity recognition and segmentation approaches to multimodal Lifelog data. In: 2019 international conference on content-based multimedia indexing (CBMI), pp 1–7

Graph and Convolution Recurrent Neural Networks for Protein-Compound Interaction Prediction

Ermal Elbasani and Jeong-Dong Kim

Abstract Machine learning-based methods play an important role in bioinformatics to predict the protein-compound interactions which is related closely for drug discovery. Modeling the properties and functions of protein sequences is an important, but challenging at the same time. This work, is presented developed method as representation learning for compounds and proteins, which integrate the representations, and develop a new protein-compound prediction approach a Graph Neural Network for compounds, and a Convolutional layer extended with Bi-directional Long Short-Term Memory Recurrent Neural Network framework for predicting component protein interaction. The convolution layer captures regulatory protein functions, while the recurrent layer captures long-term dependencies between the protein functions in order improve predictions. This approach improves considerably upon other models across several metrics. The results have outperformed other similar work by achieving 97% in accuracy in pre-diction. Furthermore, the combination of Convolution and Bi-directional LSTM tends to provide a more reliable result for the prediction.

Keywords GNN · CNN · RNN Bi-LSTM · Protein-Compound Interaction · Drug Discovery

1 Introduction

Identifying interactions between compounds and proteins is important in the discovery and development of safe and effective drugs. Drugs are usually small

E. Elbasani · J.-D. Kim (✉)
Department of Computer Science and Engineering, Sun Moon University, Asan, South Korea
e-mail: kjdvhu@gmail.com

E. Elbasani
e-mail: ermal.elbasani@gmail.com

J.-D. Kim
Genome-Based BioIT Convergence Institute, Asan, South Korea

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_13

molecular compounds which execute their actions mainly by regulating the biological functions of their targets [1]. Meanwhile, the vast majority of the targets are disease-related proteins. If a drug interacts with a protein involved in a disease, the drug can possibly be used to treat the corresponding disease [2]. Hence, the identification of compound—protein interactions is an important step in a drug discovery pipeline, which is helpful for understanding drug action mechanisms, drug side effects and disease pathology [3]. To identify potential Protein-compound effectively, machine learning-based prediction methods have been developed from a chemogenomic perspective [4], which considers the chemical space, genomic space, and its interactions in a unified framework. Graphs has played an important role as a representation learning method in many fields of data science for modeling the raw data in an easy way to be used for later computational process.

Recently, among various machine learning methods, deep neural networks (DNNs) have achieved excellent performance for various problems, such as speech recognition and visual object recognition. DNNs have been used for compounds and proteins [5, 6] such DNN- based methods do not apply end-to-end representation learning and depend on molecular fingerprints and protein family databases as in- put features, which are fixed in the DNN training process. Convolutional neural networks (CNN) and Recurrent Neural Networks (RNN) are variation of DNN that are successful methods for classifying datapoints in structured data.

In this work, we propose a combination Graphs and Convolutional and Recurrent Neural Networks for achieving a better performance in predicting interaction of protein and compounds.

The remainder of this paper is organized as follows. Section 2 present the related work. In Sect. 3 is described the proposed method and architecture following with results in Sect. 4, concluding with Sect. 5.

2 Related Work

In the new framework of modern pharmacology, the relationship between a compound and protein can be depicted as a network, in which each node represents a compound or a protein, and an edge indicates the interaction between a compound and a protein. Based on the new paradigm, many in silico network-based approaches have been introduced to predict compound—protein interactions [7–9]. These methods have some limitations, such as simulating the interaction between a compound and a protein as a bipartite network, and ignoring the similarities between compounds as well as the interactions between proteins.

Various types of compound and protein features and algorithms to predict Protein-compound have been investigated. For example, Jacob and Vert [10] used the tensor product-based features between chemical substructures and protein families and then applied SVMs with pairwise kernels. Furthermore, at [11] uses CNN with combination Graphs in order to find the protein ligand binding with a pairwise model. Also, at [12] a similar structure is used for prediction the compound interaction. A

CNN structure is used for analyzing the protein structure and GNN for the molecular structure and the vector is produced by each branch to be concatenated for the final prediction.

In our study the novel approach stays on the architecture of CNN after the max pooling layer is followed by a Bi-LSTM layer. Our rationale for including a recurrent layer after the max pooling layer is that protein functions can follow a regulatory pattern by physical constraints that dictate specific biological arrangements and frequencies of combinations of protein functions, a feature associated with specific functional elements [12, 13].

3 Materials and Method

This study describes the architecture for modeling protein sequence and molecules. Combination as pair of GNN and CRNN are selected. Dataset used to perform this task contains the smiles and protein sequence interaction, are received from the work [14]. The method that this work proposes Graph and Convolution Recurrent Neural Networks (GCRNN) contain three main components GNN CNN and RNN with Bi-directional LSTM (Fig. 1).

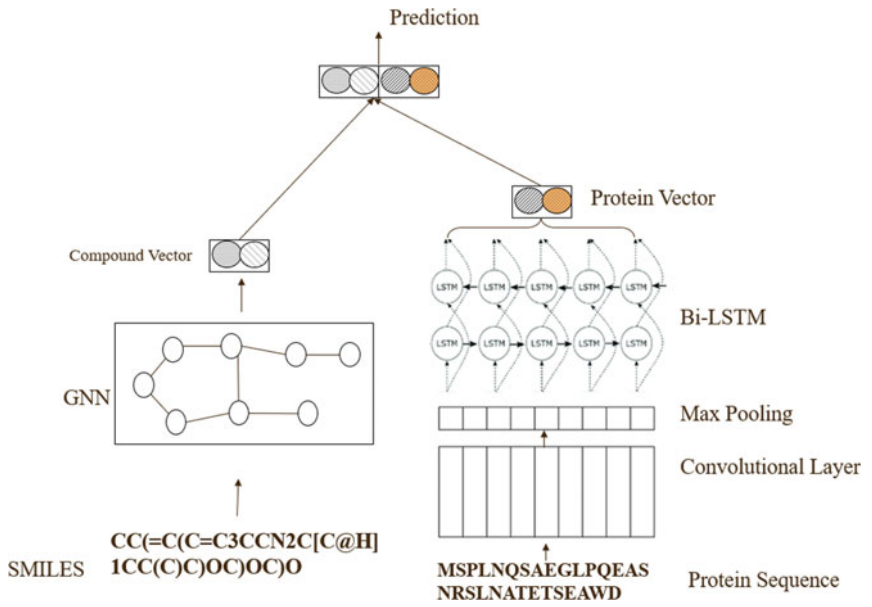


Fig. 1 Architecture of GCBI-LSTM for protein-compound prediction

3.1 Graph Neural Network for Molecular Graph

In this section, we describe the GNN, which can obtain low-dimensional real-valued vector representations of molecular graphs. A GNN maps a graph G to a vector with two functions, i.e. transition and output functions, Here, the transition function updates each vertex information in consideration of its neighboring vertices and edges in G, and the output function maps the set of vertices to the vector. Note that both functions are implemented using neural networks. This is end-to-end learning and this differentiable function requires invariance in symmetry and isomorphism for arbitrary sized, shaped and complex structured graphs each other.

3.2 Convolutional Neural Network for Protein Sequence

Convolutional neural networks (CNNs) are variants of DNNs that are appropriate for analyzing protein sequence [13, 15]. CNNs use a weight-sharing strategy to capture local patterns in data such as sequences. This weight- sharing strategy is especially useful for studying DNA because the convolution filters can capture sequence protein functions, which are short, recurring patterns in DNA that are presumed to have a biological function.

3.3 Recurrent Neural Networks

Another variation of DNNs is the recurrent neural network (RNN). Unlike a CNN, connections between units of an RNN form a directed cycle. This creates an internal state of the network that allows it to exhibit dynamic temporal or spatial behavior. A bi-directional long short-term memory network (Bi-LSTM) is a variant of the RNN that combines the outputs of two RNNs, one processing the sequence from left to right, the other one from right to left. Instead of regular hidden units, the two RNNs contain LSTM blocks, which are smart network units that can remember a value for an arbitrary length of time. BLSTMs can capture long-term dependencies and have been effective for other machine learning applications [16].

3.4 Experiments

A number of experiments have been done in order to find the differences in effectiveness between different architectures. Dataset experimental analysis is obtained from [14]. The GNN takes SMILES as input, which is a string encoding of a molecule. Note that SMILES was converted to a graph representation using RDKit. We then

extracted various information of the molecular graph, such as atom types, chemical bonds, and the adjacency list of atoms. For proteins, the pre-processing is not required because our CNN takes a raw amino acid sequence as input.

4 Results and Discussion

In this section, we analyze the effects of the Bi-LSTM after the Max pooling Layer on CNN for affecting the protein-compound interaction prediction performance. Using this method was achieved an accuracy of 97% of the data set as given in Fig. 2 and a smooth los function described in Fig. 3. In comparison with other methods form GCNN [12] our method shows a small improvement in overall performance given in Table 1. This can be considered confident achievement, since the improved

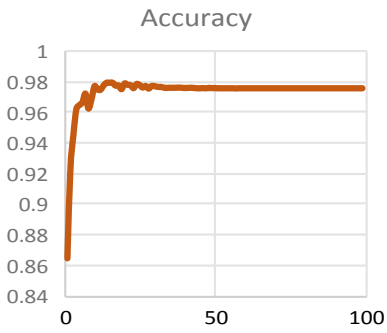


Fig. 2 The accuracy of GCRNN

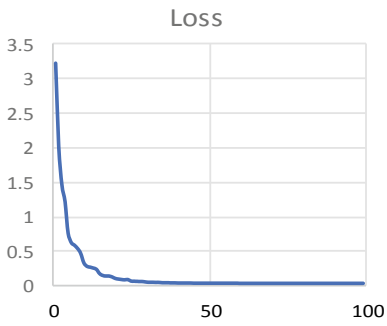


Fig. 3 Loss function of GCRNN

Table 1 Results of methods for prediction performance

Method	Precision	Recall	Accuracy
GCNN	0.92	0.94	0.96
GCRNN	0.95	0.92	0.97

method may provide a more reliable predication since protein function extractor is an important tool for interaction prediction.

5 Conclusion

This work proposes an end-to-end representation learning of a GNN and CRNN with Bi-directional LSTM to predict Protein-compound interaction. The experimental results have demonstrated that a relatively low-dimensional and shallow neural network has the potential to outperform various existing methods on both balanced and unbalanced datasets. In addition, our attention mechanism has provided clear visualizations that make real-valued vectors easier to analyze. We believe that our study will provide new insights into protein compound interaction prediction to construct a general machine learning in bioinformatics rather than using feature engineering.

Compared with traditional wet experiments, the proposed approach can predict the interactions of compound–protein interactions on accuracy at 97%. In contrast to the existing structure-based computational approaches, the current method shows good performances with only the protein primary structure information, and without three-dimensional structure information.

Acknowledgements This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019R1F1A1058394).

References

1. Li ZC, Huang MH, Zhong WQ, Liu ZQ, Xie Y, Dai Z, Zou XY (2016) Identification of drug–target interaction from interactome network with 'guilt-by-association' principle and topology features. Bioinformatics 32(7):1057–1064
2. Shi JY, Yiu SM, Li Y, Leung HC, Chin FY (2015) Predicting drug–target interaction for new drugs using enhanced similarity measures and super-target clustering. Methods 83:98–104
3. Hao M, Wang Y, Bryant SH (2016) Improved prediction of drug-target interactions using regularized least squares integrating with kernel fusion technique. Anal Chim Acta 909:41–50
4. Bredel M, Jacoby E (2004) Chemogenomics: an emerging strategy for rapid target and drug discovery. Nat Rev Genet 5:262–275
5. Hamanaka M et al (2017) Cgbvs-dnn: prediction of compound–protein interactions based on deep learning. Mol Inf 36:1600045
6. Wan F, Zeng J (2016) Deep learning with feature embedding for compound–protein interaction prediction. bioRxiv, 086033
7. Lu Y, Guo Y, Korhonen A (2017) Link prediction in drug-target interactions network using similarity indices. BMC Bioinf 18(1):39
8. Zong N, Kim H, Ngo V, Harismendy O (2017) Deep mining heterogeneous networks of biomedical linked data to predict novel drug–target associations. Bioinformatics 33(15):2337–2344
9. Nascimento AC, Prudêncio RB, Costa IG (2016) A multiple kernel learning algorithm for drug-target interaction prediction. BMC Bioinf 17(1):46

10. Jacob L, Vert J-P (2008) Protein-ligand interaction prediction: an improved chemogenomics approach. Bioinformatics 24:2149–2156
11. Jones D, Kim H, Zhang X, Zemla A, Stevenson G, Bennett WD, ... Allen JE (2020) Improved protein-ligand binding affinity prediction with structure-based deep fusion inference. arXiv preprint arXiv:2005.07704.
12. Tsubaki M, Tomii K, Sese J (2019) Compound–protein interaction prediction with end-to-end learning of neural networks for graphs and sequences. Bioinformatics 35(2):309–318
13. Zhou J, Troyanskaya OG (2015) Predicting effects of non coding variants with deep learning—based sequence model. Nat Methods 12:931–934
14. Liu H et al (2015) Improving compound–protein interaction prediction by building up highly credible negative samples. Bioinformatics 31:i221–i229
15. Quang D, Xie X (2016) DanQ: a hybrid convolutional and recurrent deep neural network for quantifying the function of DNA sequences. Nucleic Acids Res 44(11):e107–e107
16. Sundermeyer M, Alkhouli T, Wuebker J, Ney H (2014) Translation modeling with bidirectional recurrent neural networks. In: EMNLP. Doha, pp 14–25

Current Status and Forecast of Blockchain Application in Security Technology

Jiwon Han, Yerin Yu, Hanmi Lee, and Jungho Kang

Abstract The security industry has recently paid attention to distributed ledger technology as the fundamental concept of blockchain, and it is utilizing the technology's transparency and integrity features in a variety of security applications. This paper reviews the case of KT—a Korean company that implemented the blockchain-based IoT security solution platform to solve IoT security vulnerabilities—the case of Shh Korea, which applied blockchain technology to enhance the security of a messaging platform, and the case of Kakao, which applied blockchain technology to PKI for a payment authentication system. The current problems as well as the potential and limitations are then analyzed.

Keywords Blockchain · Information security

1 Introduction

As the crucial technology for cryptocurrency, which has been a major issue for the past few years, blockchain has become the key technology for the next-generation security field. Security companies are using the distributed ledger component of blockchain for the originality certification of e-documents and user authentication by guaranteeing data integrity.

According to The Truth Machine, a book written by former Wall Street Journal reporters, security is the ideal post-cryptocurrency application of blockchain technology [1]. They argued that the fundamental shift of the security paradigm was necessary since it would be difficult to solve security problems with the centralized paradigm in the era of IoT wherein all things are connected to the network. A decentralized security structure based on blockchain is suggested as the next-generation security paradigm to solve the problem.

This work was financially supported by Baewha Women's University

J. Han · Y. Yu · H. Lee · J. Kang (✉)
Department of Information Protection, Baewha Woman's University, Seoul, Korea
e-mail: hjw99909@gmail.com; kjh7548@naver.com

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_14

As such, this paper seeks to investigate blockchain technologies applicable in security and present the current status of the application and forecast of blockchain in security technology.

2 Blockchain

Blockchain is the technology of linking data units called "blocks" with "chains" and distributing and storing them at regular intervals. It is based on the decentralized P2P mode through the autonomous connections of individual nodes rather than conventional server-client mode. It is also called the Distributed Ledger Technology (DLT) since multiple network participants verify and approve the veracity of open data to secure data safely and efficiently with the need for a central authority to guarantee the data.

Since blockchain assures the CIA (confidentiality, integrity, and availability) factors, it can provide better flexibility, encryption, audit, and transparency. The distributed structure of blockchain is also relatively safe from attacks like DDoS or hacking threats since its overall performance is not affected even if a part becomes non-operational.

3 Utilization of Blockchain in Security

This section is organized as follows: 3.1 reviews the vulnerabilities of conventional IoT and the blockchain application to mitigate them; 3.2 examines the development and current utilization status of blockchain in message platforms; 3.3 investigates the utilization status of blockchain in authentication services; 3.4 reviews the problems of DNS and the blockchain application to correct them; finally, 3.5 discusses the problems of the blockchain applications above in security.

3.1 IoT Security Using Blockchain

The importance of IoT security has increased as the utilization of IoT—which is the technology of connecting a wide range of objects using sensors and communication technology—has broadened. IoT has many security vulnerabilities due to the lack of authentication mechanism and data integrity. As such, many studies are ongoing to secure the authentication mechanism and data integrity of IoT using the characteristics of blockchain, such as decentralization and P2P, to solve such IoT security problems.

Korean company KT implemented a blockchain-based IoT security solution platform that assigns a unique blockchain ID and issues a one-time token to all communication elements to solve the IoT security vulnerabilities. IoT terminals applying the platform do not expose their IP address to prevent intrusion by hackers. Moreover, even if an attacker finds the IP and attempts to intrude, the platform blocks unauthorized servers and user packets to prevent damage [2].

3.2 Safe Messaging Platform Using Blockchain

As crimes involving intrusion into information and communication networks have increased, security is likely to become the key competitiveness factor of messaging platforms. Message security usually uses end-to-end encryption, which refers to encryption technology that allows only the sending and receiving parties to read the message. Note, however, that even end-to-end encryption does not guarantee perfect security, with an increasing number of messaging platforms adopting blockchain technology to improve safety.

Shh Korea, a subsidiary of Paxnet, has applied blockchain technology to optimize user privacy protection. The blockchain mainnet developed by the company has adopted GRIDA to block hacking fundamentally. Moreover, it encrypts not only text data but also images, videos, and all data to provide more powerful security than the existing messaging platforms [3].

The "US Defense Digital Modernization Strategy" announced in 2019 includes a strategic plan to enhance the communication sector by utilizing blockchain. The US DARPA (Defense Advanced Research Projects Agency) has been sponsoring the development and testing of the hacking-proof safe messaging platform based on blockchain [4] (Fig. 1).

Since blockchain technology is based on safe, authenticated communication, its application in this area is expected to advance further.

3.3 Blockchain-Based PKI

PKI (Public Key Infrastructure) is the infrastructure that facilitates safe communication in e-commerce by having the authorized authentication agency issue and managing the user key and certificate.

Applying blockchain technology to certificate services can improve security and reliability and enable the securities firms to share the authentication information. It also allows fast transactions through P2P technology since it enables the distributed storage of transaction records. More studies have recently reported the application of blockchain technology in PKI.

Korean company Kakao introduced the activeX-free, blockchain-based PKI system for payment authentication services by utilizing the advantage of blockchain.

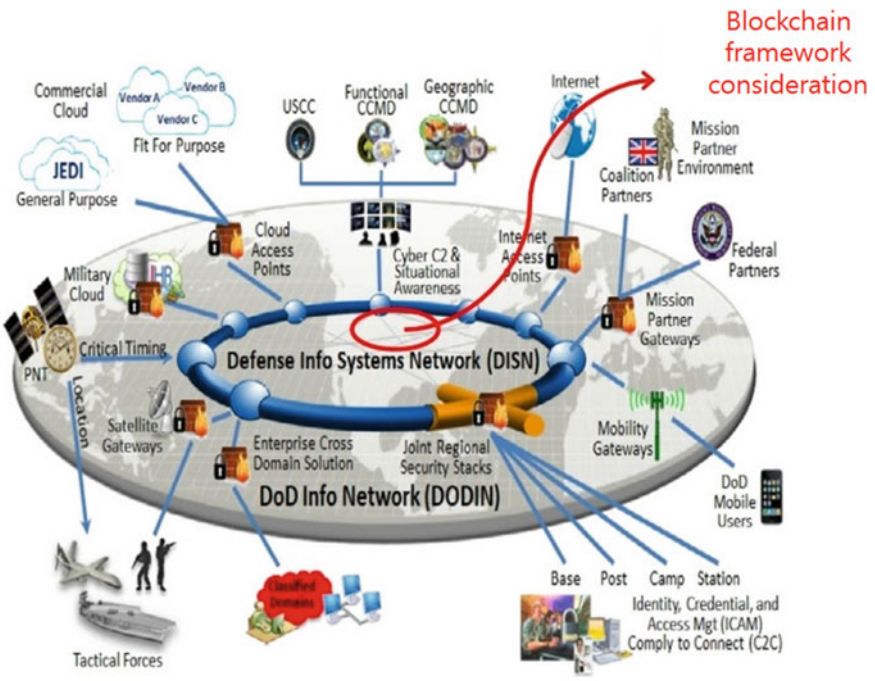


Fig. 1 "DoD Digital Modernization Strategy," DoD Office Prepublication and Security Review, 2019.7.12

The existing activeX-based public-key cryptography has had the problem of secret key management that is vulnerable to security. The company implemented an electronic signature platform using blockchain to solve the problem. Although initially using the bitcoin blockchain, it currently utilizes the hyperledger fabric since bitcoin blockchain has a limitation in storing certificates [5].

3.4 Improvement and Expansion of DNS Using Blockchain

The current domain, which plays an important role in mapping many hosts and their addresses, has several problems. First are the complex registration procedure and comparison and analysis of many proxies for users to register and use the domain name of the DNS. Second are the shortage of domain names and intensifying disputes concerning domain name registration. The last problem involves security vulnerabilities [6]. To overcome such problems of DNS, there are many studies of expanding and improving DNS through decentralization using the blockchain network, such as namecoin and EmerDNS (Fig. 2).

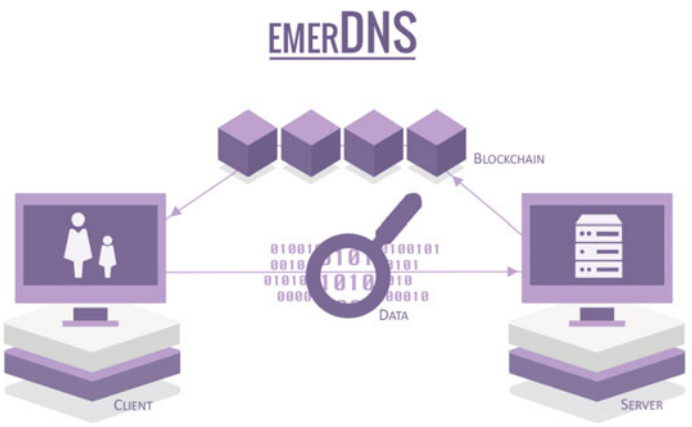


Fig. 2 "Emerdns," Emercoin

Namecoin is the leading example of DNS applying blockchain. It is a new-concept DNS that applies blockchain technology and is used by the bitcoin system, a type of cryptocurrency that aims to decentralize DNS and ID. The namecoin system was redeveloped from the bitcoin system, and the additional key value contains the domain name and IP address. Nonetheless, the namecoin has failed to gain the attention of users, and the exchanges no longer trade it.

3.5 Problems of Blockchain Applications

As described in this paper, blockchain is used by the information security field to solve existing problems by taking advantage of blockchain features such as decentralization and integrity. Like the cases described above, however, blockchain applications to solve the security problems have also shown vulnerabilities.

First are the problems in IoT. The blockchain applied to IoT inspects all blocks to ensure integrity as the strength of blockchain. Note, however, that the inspection time increases proportionally to the increasing data volume; thus, the service speed slows down when the data volume is large. Moreover, while blockchain enhances the security of IoT networks, it is not perfect because it does not guarantee the security of devices [7].

Second are the problems in PKI. Many blockchain applications in PKI are being developed because it has proven to be very useful. Still, it has limitations in PKI services because of the long time required for inspection when recording the data.

The last problem is DNS. The namecoin was a successful application of blockchain in DNS but failed to gain users' attention due to the inconvenience of having to install the program in the web browser to log in to the ".bit" site through the namecoin DNS information and the high use fee [8] (Table 1).

Table 1 Security vulnerabilities of blockchain

Area	Security Vulnerabilities of Blockchain
IoT	* Increased inspection time, service slowed down as a result * No guaranteed device security despite the improved security in the IoT networks
PKI	* Service slowed down by the long inspection time for data registration
DNS	* Inconvenience of having to install the program in the browser

4 Forecast of Blockchain in Security

This paper reviews the application of blockchain in security fields and their problems. As described in Sect. 3.5, there are additional problems of blockchain applications that cannot be solved with the current technology [9]. Nonetheless, blockchain applications in security will certainly continue to increase. This is because blockchain technology plays a significant role in supplementing and improving many problems in security, and it is highly likely to advance further in the era of the Fourth Industrial Revolution wherein IoT is crucial.

The academe and industries are applying to information security technology blockchain that cannot be tampered with or falsified, and global companies are investing more in blockchain technologies [10]. Moreover, blockchain is expected to lead the expansion of the big data market by enhancing not only data security but also personal control of individual data [11].

5 Conclusion

This study sought to survey the current status of blockchain in information security and to study its utilization and forecast in the future.

As pointed out by this paper, blockchain is solving many problems in security. Blockchain not only reduces risks such as information leak and damage but also provides benefits and solutions to problems, such as ensuring integrity and providing fast and transparent transaction details through horizontal linkage. Nonetheless, it also gives rise to other problems as side effects.

Most of the problems of blockchain application in security are related to service deficiency due to the prolonged block inspection time. It is a problem that occurs as the blockchain inspects blocks to ensure integrity during the data recording and storage process. Solving the problem requires seeking improvement measures to reduce block inspection time while maintaining integrity as the advantage of blockchain.

The threats of attacks such as hacking are increasing as the fourth industry develops. It is necessary to implement better security systems using blockchain to be better prepared for the threats.

References

1. Casey MJ, Vigna P, The Truth Machine: The Blockchain and the Future of Everything, translated by Kim Ji-yeon, Mirae Book, pp 195–204
2. Hwang C-G (2019) KT, Blockchain-based IoT Security Giga Stealth Pilot Service. https://www.boannews.com/media/view.asp?idx=83382
3. Min H-K, PaxNet, Blockchain-based Messenger, Shh, Open beta service. http://www.newscu.net/news/articleView.html?idxno=10157
4. Lee K-H, Park H-S (2020) Study on Trend of Defense Blockchain Technology and Defense ICT Convergent Strategy. ETRI 35(1)
5. Lee H-W (2018) ActiveX-free Blockchain-based PKI System, if(kakao) (2018)
6. Heo J-W, Kim J-H, Jeon M-S (2019) Design and implementation of Blockchain network based domain name system, pp 37–38. Korea Academia-Industrial cooperation Society
7. Lee B-M, Kang H-S, Yu S-H, Lee Y-S (2018) Security vulnerability analysis of Blockchain-based IoT service, pp 383–384. Proceedings of KIIT Conference (2018)
8. Rothstein A (2018) The end of money: the story of bitcoin, cryptocurrencies and the blockchain revolution. sciencebooks. Seoul
9. Kim H-Y (2018) Analysis of security Threats to Blockchain platform and response measures, pp 108–109. Korean Institute of Information Technology
10. Jung K-S, Kim D-W (2019) Blockchain industry status and overseas policy trend, p 14 National IT Industry Promotion Agency
11. Lee J-Y, Woo C-W (2018) Forecast, limitation, and implication of Blockchain technology, pp 14–15. Science and Technology Policy Institute

Blockchain-Based Secure Digital Twin Framework for Smart Healthy City

Abir EL Azzaoui, Tae Woo Kim, Vincenzo Loia, and Jong Hyuk Park

Abstract Nowadays, Digital Twins (DTs) are being integrated into various sectors thanks to the continuous progress of computing power and data science. We are surrounded by a tremendous number of sensors and connected objects that produce data regularly. These data represent the fuel for a DT as they are used to represent the most accurate digital model of a system or an object and to predict and simulate all the possible scenarios. Recently, DT has been adapted into the healthcare sector as well for an accurate medical and surgery simulation and medical resources' orchestration. However, DT technology is still a novice to the healthcare system and security threats urge immediate consideration. To this end, we propose in this paper a Blockchain-based secure Digital Twin framework for a smart healthy city. We discuss as a case study the current COVID-19 pandemic and argue on the help of DT to control the situation, prevent future cases, and personalize the treatment.

Keywords Digital twin · Smart healthcare · Blockchain · Security · And privacy

A. EL Azzaoui · T. W. Kim · J. H. Park (✉)
Department of Computer Science
and Engineering, Seoul National University of Science and Technology, Seoul 01811, South Korea
e-mail: jhpark1@seoultech.ac.kr

A. EL Azzaoui
e-mail: abir.el@seoultech.ac.kr

T. W. Kim
e-mail: tang_kim@seoultech.ac.kr

V. Loia
Department of Management and Innovation Systems, University of Salerno, Fisciano, Italy
e-mail: loia@unisa.it

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes
in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_15

1 Introduction

Physical models have been continuously developed in recent years. However, thanks to the advancements in computer science and the unceasing evolution in the Industrial Internet of Things (IIoT), the classical physical systems are now shifted to digital models, which is called scientifically by Digital Twins (DT). DT is a real-time data-informed virtual digital replica of a physical device or a complex system, used to model test scenarios before applying them to the physical entity. The term of DT was proposed in 2003 [1, 2] and was adopted mainly for industry 4.0 to model and simulate complex physical entities such as airplanes or even rockets. Nowadays, many researchers have studied DT for various sectors including Health care. A DT in the Health care system can be defined as the digital representation of a person, medical resources, and hospitals at large. A doctor can practice his upcoming heart surgery on the DT of his patient and simulate on it all the possible scenarios to choose the most suitable one for the condition of the patient. A hospital can have its DT system to help to manage and orchestrating human resources such as medical staff and physical resources such as medical machines and beds. DT can transform 1traditional and smart healthcare systems by reducing the time, cost, and provide more personalized treatment for the patients. Liu et al. [3] proposed a framework of the cloud healthcare system based on digital twin healthcare named (CloudDTH). The proposal aims to achieve interaction and convergence between medical physical and virtual spaces using the concept of DT. Rivera et al. [4] discussed the importance of DT for a personalized healthcare system and treatment and elaborated on the definition of internal structures for DT to support precision medicine techniques in the context of continuous monitoring and personalized data-driven medical treatments. Jimenez et al. [5] argued om the necessity to integrate DT into the healthcare sector to enhance their capabilities and offer better solutions, they provide as well some definitions of Medical Cyber-Physical Systems (MCPSs) and Digital Twins along with technological enablers such as cloud and IoT.

DT is indeed a solution for a smart and more precise healthcare system. However, it triggers various security threats. A DT urges continuous data uploading to come out with the best results. In the case of the healthcare sector, these data engender patient's personal information and medical history. Such information needs to be secured, non-falsified, and should never be leaked. Thus, in this paper, we propose the use of Blockchain to secure the user's identity and make the data available anonymously only to healthcare providers and professionals for real-time data analytics, medical researches, and personalized treatments. Using Blockchain-based DT, we can create a secure and anonymous database shared between healthcare providers to improve the accuracy of treatments, predict future diseases, and control them.

The contribution of this paper is as follows:

- We present a secure Blockchain-based Digital Twin framework for smart healthy cities.

- We propose as a case study the application of this framework on COVID-19 pandemic to discuss its usability and integrity.
- We explain the scenarios of the framework using a detailed methodology flow.

In Sect. 2, we present an overview of the proposed framework. We discuss the scenarios and sequence diagram in Sect. 3 and conclude this paper in Sect. 4.

2 Proposed Framework

In this paper, we propose a secure Blockchain-based Digital Twin framework for the future of the smart healthy city. As a case study, we decided to discuss how Blockchain-based Digital Twin can help to control the current pandemic of COVID-19. The proposed framework shown in Fig. 1 is composed of three layers: (1) Device Layer, (2) Blockchain Layer and (3) Application Layer.

The Device Layer: This layer includes the Digital Twins of persons (in this case the DT is the user's Smartphone, as it is cheap, everywhere and easy to use by a regular person) and hospitals. The physical entity (person) is responsible for feeding its respective DT (Smartphone) with accurate data. The hospitals participate as well by sending data regarding medical resources to the Blockchain Layer.

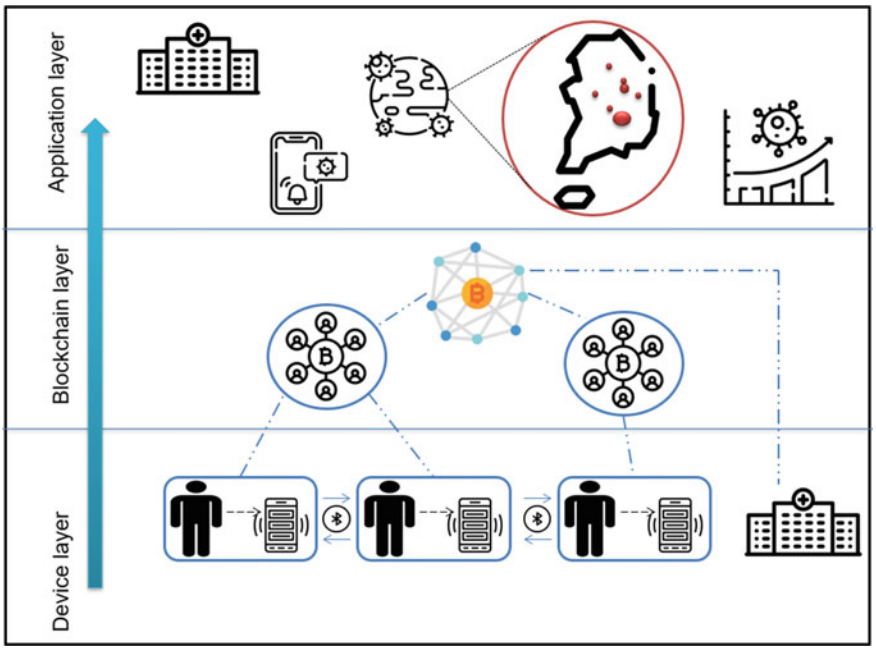


Fig. 1 Blockchain-based digital twin framework for Smart Healthy City

The Blockchain Layer: This layer has two main components, Consortium Blockchain, and Private Blockchain. Consortium Blockchain is controlled by the city's hospitals and are responsible of validating the accuracy of test results based on the giving test identification, in this context, if a person undertake a COVID-19 test in the hospital, the hospital associate his test with a unique identifier and use it to check if the data sent by the DT is accurate (if the person has indeed take the test and if the result is positive). Consortium Blockchain assures the accuracy and non-falsify of data as it is controlled by a group of approved entities [6]. Besides the Consortium Blockchain, this Blockchain layer includes as well a Global Blockchain, this Blockchain is Private and managed by one trusted organization (World Health Organization for example), however, it is visible for all the health institutions all over the world. The data collected in Blockchain are sent to the application layer. This method creates an openly shared database that can serve for future researches for a better understanding of the virus and more precise medication.

The Application Layer: Based on the data sent by the Blockchain layer, the map is updated in the application layer to show new patients' locations securely, a notification is sent to all the DT who have been around the patient's DT (based on the public keys, we explain this method on the third section) and the user's symptoms are being analyzed for better understanding and prediction.

3 Sequence Diagram

To understand how this framework works, we present a sequence scenarios' diagram as shown in Fig. 2. For clarification and simplicity, we consider in this scenario only two DTs and one hospital. The public key exchange is via Bluetooth and we consider a scenario where every smartphone has a secure secret key associated with a public key used as DT identification. The encryption and public/private key generation algorithms fall out of the scope of this study.

Step 1: The user is responsible for his own Digital Twin' data uploading, these data are divided between static data that can be added once to the user's DT such as user name, age, gender, most frequent locations such as home and work or school, and some of the user's underlying health issues. And modular data that can be changed based on the user's current symptoms; these data include Boolean data type (yes or no) that covers the usual COVID-19 symptoms such as:

- Losing sense of smell or taste.
- Trouble breathing.
- Pain in through and/or chest.
- Dry cough.
- Extreme fatigue.

A float data type for the temperature, date and time data type for the timestamp to track the change of the symptoms, a Boolean type for COVID-19 test (0 for a

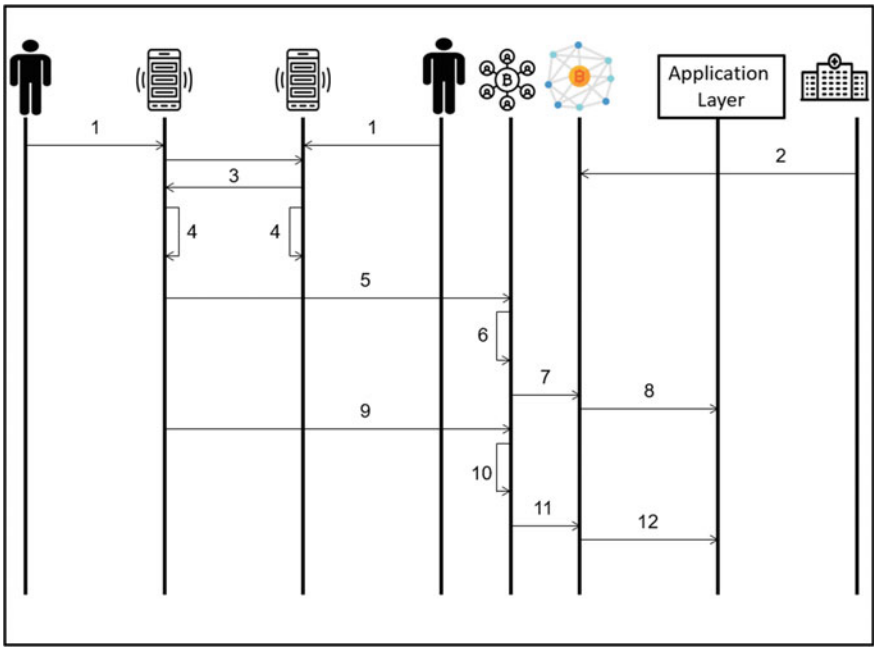


Fig. 2 Sequence diagram for the proposed framework

negative test and 1 for a positive test), and an integer data type for test identification in case a user undertook a test. All the above-mentioned information is being kept in the DT.

Step 2: Every hospital's DT records information about the available resources such as beds, medical staff, and medical equipment. These data are sent to the Global Blockchain for better resource management between deferent hospitals (i.e.: patients, based on the gravity of their symptoms, are assigned automatically to the right hospital that can respond to their needs or the hospital with enough medical resources.) thus reducing the time and cost.

Step 3: If two users or more stayed at one place for over 5 min, their DTs exchange the public keys and save them.

Step 4: The DT keeps the record of each public key for 15 days and then deletes it.

Step 5: In the case of a user being tested, and the test is positive (Test_Result == 1), the user has to type his test identification (Giving by the hospital). The DT by it turns to send the frequently visited locations' data, symptoms data, and test identification to the consortium Blockchain.

Step 6: Consortium Blockchain participants check if the test results are accurate.

Step 7: After confirming the results, the information is sent to the Global Blockchain.

Step 8: The location of the user anonymously is added as a red spot in the map, and the respective symptoms are being shared anonymously as well for the globally shared real-time database. And based on the patient underlying health conditions and current symptoms, personalized treatment is offered. Moreover, using the stored Public Key, an alert is sent to all the DTs with the respective Public Keys so they can be aware of the situation and undertake the test as well before they spread the virus more.

Step 9: The user with a previous positive test has to undertake the test again. If the result is negative, the information is sent with the test identification again to the consortium Blockchain.

Step 10: The consortium Blockchain participants check if the test results are accurate.

Step 11: After confirming the results, the information is sent to the Global Blockchain.

Step 12: The map is updated.

These steps are continuously repeated, the user must update the data based on his current symptoms and the hospital must update the data based on its current situation as well. The data shared in the Blockchain can be used for future predictions and reduce the number of future cases by showing a real-time map to avoid infected places and sending a personalized notification to the protentional patient to decrease the possibility of spreading the virus. It can be used as well as a collaborative researches platform between hospitals and healthcare providers all over the world and to create a personalized treatment for the user based on his health condition. This framework is not only designed for the COVID-19 pandemic but can also be used for future virus crises.

4 Conclusion

In this paper, we proposed a Blockchain-based Digital Twin framework for a smart healthy city. We discussed the application of this framework for the current COVID-19 pandemic and explained its sequence diagram based on a methodology flow. Digital Twins can remarkably benefit and improve smart healthcare systems by digitally modeling objects, systems as well as humans, testing all the possible scenarios, and orchestrating medical resources. In the case of the COVID-19 scenario, a DT is capable of creating personalized treatment, managing medical resources, and providing a secure shared database for healthcare providers and professionals to anonymously use the data to improve the treatment and prevent futures virus-waves from occurring.

Acknowledgements This study was supported by the Advanced Research Project funded by the SeoulTech(Seoul National University of Science and Technology).

References

1. Grieves M (2015) Digital twin: manufacturing excellence through virtual factory replication, pp 1–7
2. Grieves M, Vickers J (2017) Digital twin: Mitigating unpredictable undesirable emergent behavior in complex systems, in Transdisciplinary Perspectives on Complex Systems, Berlin, Germany: Springer, pp 85–113
3. Liu Y et al (2019) A novel cloud-based framework for the elderly healthcare services using digital twin. In: IEEE Access 7:49088–49101. https://doi.org/10.1109/access.2019.2909828
4. Rivera LF, Jiménez M, Angara P, Villegas NM, Tamura G, Müller HA (2019) Towards continuous monitoring in personalized healthcare through digital twins. In: Proceedings of the 29th annual international conference on computer science and software engineering (CASCON '19). IBM Corp., USA, pp 329–335
5. Jimenez JI, Jahankhani H, Kendzierskyj S (2020) Health care in the cyberspace: medical cyber-physical system and digital twin challenges. In: Farsi M, Daneshkhah A, Hosseinian-Far A, Jahankhani H (eds) Digital twin technologies and smart cities. Internet of Things (Technology, Communications and Computing). Springer, Cham
6. Lee Y, Rathore S, Park JH et al (2020) A blockchain-based smart home gateway architecture for preventing data forgery. Hum Cent Comput Inf Sci 10:9. https://doi.org/10.1186/s13673-020-0214-5

Blended Learning as an Effective Approach to English Language Teaching at the Institutions of Higher Learning—A Case Study

Blanka Klimova

Abstract Research shows that the blended learning (BL) is now a common teaching approach at the institutions of higher education thanks to a combination of different teaching approaches, which can satisfy the needs of both full-time and part-time students. The purpose of this article is to describe the effectiveness of the BL approach in teaching English as a foreign language (EFL) to part-time students in their third year of study at the Faculty of Informatics and Management of the University of Hradec Kralove, Czech Republic. The methods used include a literature review of available studies found on the research topic in two acknowledged databases (Web of Science and Scopus) and an evaluation of students' final credit test results and their answers to two open-ended questions. The results of this study confirm that the BL approach is effective since students were satisfied with both the face-to-face sessions and online classes. In fact, the success of teaching and learning depends on a variety of methods used, which are targeted at meeting students' needs. In this study, the BL approach was aimed at students' language, working and family needs in order to enhance their language skills, as well as enable them to fulfil their job and family commitments.

Keywords Blended learning · EFL · University students · Approach · Effectiveness

1 Introduction

There exists a number of definitions of blended learning (BL). In fact, it can be said that the implementation of BL offers teachers to choose from a variety of teaching approaches [1]. For example, Driscoll [2] perceives it as a combination of different web technologies to achieve educational goals. It can be a combination of various pedagogical approaches ensuring the optimal learning results with or without use of technologies. It can be also seen as a mixture of any type of educational technology

B. Klimova (✉)
Faculty of Informatics and Management, University of Hradec Kralove, Rokitanského 62, 500 03 Hradec Kralove, Czech Republic
e-mail: blanka.klimova@uhk.cz

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_17

with frontal lecturing. However, most often, BL is understood as a combination of traditional, face-to-face teaching and online learning [3]. Research shows [4–7] that blended learning (BL) is now a common teaching approach at the institutions of higher education, especially preferred by part-time students who come to study to university on Friday's afternoons and Saturday's mornings, at least in the Czech Republic.

In addition, the findings of the research studies reveal that the BL approach is effective [8–12]. For example, Rios and Cabrera [8] report that their students in the experimental group, which employed the BL approach, attained a significant language improvement compared to the control group that used a traditional face-to-face method. Furthermore, the findings of a Chinese case study [9] also reveal the effectiveness of the BL approach. This was evidenced on the basis of the results obtained from a pretest, progress test and post-test, respectively, which showed that the students in the experimental class had achieved higher scores in their post-test at the end of the study.

In addition, as Kachmarchyk et al. [13] indicate, the BL approach ensures that the key educational process requirements (openness, accessibility, interactivity, and instrumentality) positively affects both teachers and students. Furthermore, it improves learners' interaction, raises their motivation and provides autonomous learning [14]. As Prihastiwi et al. [1] point out, learning autonomy is vital in the context of teaching English as a foreign language. They state that without learning autonomy, learners study in limitations. For all these reasons described above, the BL approach appears to be a very suitable method for English language teaching at a university level.

The purpose of this article is to discuss the effectiveness of the BL approach in teaching English as a foreign language (EFL) to part-time students in their third year of study at the Faculty of Informatics and Management of the University of Hradec Kralove, Czech Republic.

2 Methods

2.1 Methodology

Firstly, a literature review of available studies found on the research topic, i.e., *blended learning* AND *teaching English as a foreign language*, was performed in two databases: Scopus and Web of Science. Altogether 178 articles were detected in both databases; the oldest dating back to the year of 2005. In addition, 11 articles were identified through other sources and from the references of the detected database studies. After a thorough selection process, only 23 studies remained for the final analysis and they are listed in the part on References.

Secondly, the effectiveness of the BL approach was investigated on the evaluation of students' final credit test results and their answers to two open-ended questions:

1. Did you find learning materials in the online course useful? Why yes or not? 2. What did you miss both in the face-to-face sessions and online course?

Thirdly, since the subject sample was small, a method of case study was used to describe the research question, which is as follows:

Is the BL approach effective in teaching English as a foreign language at a university level?

2.2 Study Design

Altogether 21 part-time students of Management of Tourism in their third year of their study at the Faculty of Informatics and Management attended the Course of Practical English language in the winter semester of 2019. The semester lasted for 13 weeks. However the part-time students had only two face-to-face contact sessions per this semester, each consisting of six 45 min lessons. Otherwise they had to study on their own by using study materials from the online blackboard course (Fig. 1). During the contact sessions, the teacher explained to and practiced with the students the study materials contained in this online course. The final credit test was based only on the materials in the online course. All the exercises which had to be filled in or translated into English were complemented with the key that contained the right answers.

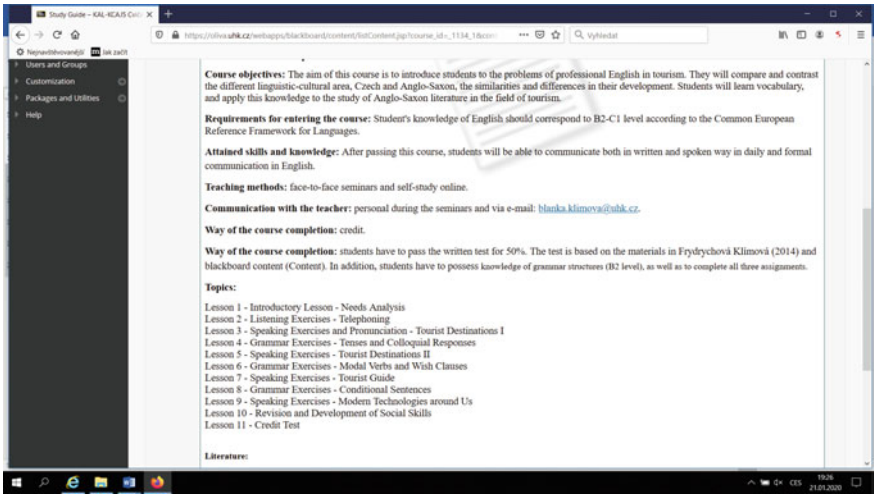


Fig. 1 An overview of the online Course on Practical English Language [15]

3 Results and Discussion

The findings of the final credit test, which had to be passed for 50%, i.e. 20 points, and which was taken by all students (21), showed that only three students had not passed this test. However, these three students did not attend the face-to-face sessions and therefore, they did not have a chance to listen to teacher's explanations and also practice the language skills through the exercises contained in the online course.

As far as the evaluation questions are concerned, students were very satisfied with the learning materials in the online course since they had everything in one place, the structure was easy to follow and the learning materials easy to find. Moreover, the students found these materials useful not only for the preparation for the final test, but also for the enhancement of their language skills. The students also appreciated the key with the correct answers to all exercises so that they could practice them on their own as well. On the contrary, students would welcome more face-to-face classes during which they would be able to practice speaking skills more. However, this is a drawback of the part-time studies.

Overall, it seems that the BL approach is effective since students were satisfied with both the face-to-face sessions and online classes. In addition, students succeeded in passing the final credit test and felt that they had improved their language skills. Therefore, the BL approach is effective because students were not only satisfied, but they were motivated to study on their own, performed well and thus were able to construct new knowledge [16]. The effectiveness of BL was then influenced by design features of BL, e.g. technology quality, interactions, LMS tools and resources and face-to-face support, as well as by learner characteristics, such as self-regulation, attitude, computer competence, age, or social support [16].

For further implications connected to the described research see the connected research into the use of mobile technologies in learning foreign languages and various aspects of on-line learning methodologies reflected in intercultural communication and business communication [17–22].

4 Conclusion

The results of this study confirm that the BL approach is effective since students were satisfied with both the face-to-face sessions and online classes. In fact, the success of teaching and learning depends on a variety of methods used, which are targeted at meeting students' needs. In this study, the BL approach aimed at students' language, working and family needs in order to enhance their language skills, as well as enable them to fulfil their job commitments.

Acknowledgements This article is supported by the SPEV project 2020, run at the Faculty of Informatics and Management, University of Hradec Kralove, Czech Republic. The author thanks Aleš Berger for his help with the data collection.

References

1. Prihastiwi L, Rochsantiningsih D, Suparno: Blended learning to activate English as a foreign language (EFL) learners' autonomy. Pertanika J Soc Sci Humanit 25:1–9
2. Driscoll M (2003) Blended learning: Let´s get beyond the hype. https://www-07.ibm.com/services/pdf/blended_learning.pdf
3. Frydrychova Klimova B (2009) Blended learning. In: Mendez Vilas A et al. (Eds), Research, reflections and innovations in integrating ICT in education pp 705–708. Spain: FORMATEX
4. Frydrychova Klimova B, Poulova P (2014) Forms of instructions and students' preferences—a comparative study. In: Proceedings of the 7th international conference (ICHL 2014), pp 220–231. Heidelberg: Springer
5. Buran A, Evseeva A (2015) Prospects of blended learning implementation at technical University. Procedia Soc Behav Sci 206:177–182
6. Simonova I, Kostolanyova K (2016) The blended learning concept: comparative study of two universities. In: Blended learning: aligning theory with practices. LNCS. 9757, 302–311
7. Cerna M (2016) User evaluation of language websites as a way of students' engagement into blended learning process case study. In: Blended learning: aligning theory with practices. LNCS 9757, 269–280
8. Rios SM, Cabrera AF (2008) Effectiveness of a blended learning model for English as a foreign language (EFL) teaching: an empirical study. Rla-Revista de Linguistica Teorica y Aplicada 46(2):95–118
9. Sun ZN, Qiu X (2017) Developing a blended learning model in an EFL class. Int J Contin Eng Educ Life-Long Learn 27:1–2
10. Nikolaeva S, Zadorozhna I, Datskiv O (2019) Development of pre-service English teachers' language skills and learner autonomy via blended learning. Revista Romaneasca Pentru Educatie Multidimensionala 11(2):222–239
11. Klímová B, Pražák P (2019) Mobile blended learning and evaluation of its effectiveness on students' learning achievement. In: Cheung S, Lee LK, Simonova I, Kozel T, Kwok LF (eds), Blended learning: educational innovation for personalized learning. ICBL 2019. Lecture Notes in Computer Science 11546, 216–224
12. Al Bataineh KB, Banikalef AEAA, Albashtawi AH (2019) The effect of blended learning on EFL students' grammar performance and attitudes: an Investigatin of moodle. Arab World English J 10(1):324–334
13. Kachmarchyk S, Khrystiuk S, Shanaieva-Tsymbal L (2019) Using blended learning technology in foreign language communicative competence forming of future international relations specialists. Revista Romaneasca Pentru Educatie Multidimensionala 11(4):84–99
14. Bukhari SSF, Basaffar FM (2019) EFL learners' perception about integrating blended learning in ELT. Arab World English J 5:190–205
15. Course of Practical English Language. https://oliva.uhk.cz/webapps/blackboard/content/listContent.jsp?course_id=_1134_1&content_id=_143670_1&mode=reset (2020)
16. Kintu, M.J., Zhu, C., Kagambe, E.: Blended learning effectiveness: the relationship between student characteristics, design features and outcomes. Int J Educ Technol High Educ 14, 7 (2017)
17. Pikhart M (2014) New horizons of intercultural communication: applied linguistics approach. Procedia Soc Behav Sci 152:954–959
18. Pikhart M (2018) Electronic managerial communication: new trends of intercultural business communication. Innovation management and education excellence through vision 2020, I -XI, 714–717
19. Pikhart M (2018) Communication based models of information transfer in modern management—the use of mobile technologies in company communication. Innovation management and education excellence through vision 2020, I -XI, 447–450
20. Pikhart M, Klímová B (2019) Utilization of linguistic aspects of bloom's taxonomy in blended learning. Educ Sci 9(3):235

21. Klímová B, Pikhart M (2019) Mobile applications and their role in teaching english—methodological basis. ICERI2019 Proceedings 2th annual international conference of education, research and innovation, pp 986–990
22. Pikhart M, Klímová B (2019) Digital transformation of intercultural communication. ICERI2019 Proceedings 2th annual international conference of education, research and innovation, pp 991–994

Cultural Discourse in Human-Computer Interactions: Linguistic Aspects of Digital Communication

Marcel Pikhart

Abstract Human-computer interaction (HCI) has drawn a lot of attention in the past few years from the technological viewpoint, however, there are several aspects that must be considered when dealing with optimizing HCI. Information technology is ubiquitous and so is digital communication which is conducted globally through the network, both the Internet and intranets of companies. Digital communication bears a significant aspect, similarly to any kind of communication, and it is the fact that it is based on cultural discourse. Due to the fact that any digital communication and human-computer interaction are culturally conditioned and it takes place in the global world of interconnectedness, it is always intercultural. The cultural discourse, manifested by linguistic means, of this communication, is one of the most important parameters which must be taken into account if we want to optimize information transfers and attempt to reach relatively lossless information transfer. The paper summarizes the results of the research which dealt with the issue of cultural narratives present in websites. The findings clearly show that the ICT students do not realize the importance of this aspect despite the fact that in business and other related areas the cultural narratives have already been researched sufficiently. The paper will be useful for ICT departments and managers as it draws our attention to this crucial aspect of HCI necessary for future technology enhancements.

Keywords Human-computer interaction · HCI · Communication · Intercultural communication · Cognitive science · Cultural narratives · Cultural discourse · Cultural conditioning · ICT and culture

M. Pikhart (✉)
Faculty of Informatics and Management, University of Hradec Kralove, Hradec Kralove, Czech Republic
e-mail: marcel.pikhart@uhk.cz

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_18

1 Introduction

1.1 Literature Review

Human-computer interaction (HCI) has drawn a lot of attention in the past few years due to ubiquitous computing, Internet of things, artificial intelligence, deep learning, machine learning, cognitive science and all related aspects of our digital world [1–3], however, there is an aspect which plays a crucial part in this interaction, which is a cultural narrative present basically in all aspects of HCI [4–6]. This aspect has already been described well in business studies and management theory [7–9], however, computer science still lacks behind the current research into the area [10–14].

Human-computer interaction is, as the title suggests, a bipolar encounter, i.e. a human vs a computer. This bears a lot of important parameters that must be taken into consideration when trying to improve any kind of HCI. The digital nature of computers interacts with the cultural nature of humans. Culture is defined as the mental programming of the human mind [5, 6], i.e. it is the operating system upon which the humans work, cooperate, negotiate and deal with each other.

Human-computer interaction is also present in various other areas, such as eLearning, mLearning, blended learning and hybrid learning, where the importance of the culturally specific environments is even enhanced, and this area has been researched widely enough [15–23]. This research is sufficient and does not need a further extension.

ICT departments and all its related functions still lack the proper approach which could improve digital communication and HCI significantly.

2 Research

2.1 Research Design

The research was conducted at the Faculty of Informatics and Management during the winter semester of the academic year 2019/2020 in a group of IT students into their opinions on how much it is important to take into account cultural discourse when creating websites for various companies. They were also researched into their realization of cultural aspects of web design in various websites they have seen and used in their professional life as templates or inspiration.

The data were collected by an online questionnaire distributed to them at the end of the semester. The number of respondents was 23, both male and female, aged 22–24, all of them were Czech nationals.

2.2 Research Question

The research question was: How much is cultural discourse important when implementing it into websites which you create?

The paper claims that the cultural discourse implemented into websites is crucial for understanding the message the website conveys. Therefore, this research question was used in the respondents.

2.3 Research Findings

The research clearly shows that the relevance interculturality, cultural discourse and cultural narratives are not considered to be crucial for the respondents. The research question was put simply with another set of sub-questions which should have fine-tuned the respondents´ viewpoint and ideas. Their concept is merely technical and technological when designing their websites so that there is not much space for any significant cultural aspect. However, after giving them the questionnaire and conducting subsequent interviews, the respondents replied in a very positive way to the information about the relevance of cultural discourse for website design.

The most important finding is that the respondents do not realize any relevance of culturally biased narratives in their websites. This could create potential threats for information transfer through ICT. A practical example could be misunderstood contents of the website due to an inappropriate layout and the use of colours in various cultural environments where colour can have a negative connotation in one culture and a positive one in another. Eye-tracking research will also bring very many important findings rooted culturally, such as the trajectory by which we scan the website in different cultures, based on the writing style of the given region (Asia vs. Europe).

The respondents claimed that there had not been any systematic training into these aspects of culture and they themselves called for an action, i.e. to improve the university curricula, so that they abandon their technicality but would be a little bit closer to important and burning issues of the current global world of information transfer, in the context of artificial intelligence, internet of things, biotech and infotech. Philosophy, ethics and related disciplines will soon become crucial in the context of the global ICT world as they give crucial insight which must be combined with technical and engineering aspects of data processing.

2.4 Research Limitations

The research was conducted with a small group of respondents. They are not very experienced programmers but many of them have some relevant programming experience from companies where they have a part-time job. Despite all these concerns, we can claim that there is significant relevance of the research results and we can assume that very similar results would be obtained with a much larger sample, based on the personal experience of the researcher.

3 Discussion

The findings of the research, when put in the context of previous research [10, 12, 15], clearly show that the awareness of the topic is still very low, not only in IT students but also in IT professionals. The situation is much better in business communication and management studies which have already widely accepted these realities of the global world and global communication and count on the importance of interculturality, cultural narratives and cultural discourse. These cultural aspects are very much related to the linguistic aspects of our discourse, i.e. the language we use when we communicate.

ICT departments should draw on findings which we have already available from business research and interculturality in business and managerial communication [6, 7, 9] as these findings can significantly improve the basic setting of the information and communication technology procedures. When comparing the results of this research in business, we can conclude that the findings are basically unanimous, therefore, the paper claims they can easily be transferred into the area of ICT expertise. Business and ICT deal hand in hand with communication issues and it is the ICT department that should be liable for any information transfer deficiencies. The realizing of the significance of these cultural aspects will present a step forward towards enhanced digital communication in the global world.

Further research into this topic could be useful to obtain more accurate data about the area of the importance of cultural narratives and discourse present in websites of global companies so that the smaller ICT companies could have more inspirational ideas when designing their own websites.

4 Conclusion

This paper attempts to highlight the importance of cultural narratives which are present in human-computer interactions because they are often neglected even if they present a threat to the lossless transfer of information if they are ignored.

The contribution of the paper is to draw attention to the importance of intercultural discourse present in websites, the idea which is still relatively novel in ICT departments. The findings are only preliminary but they will be crucial for ICT departments that are responsible for the information systems used by the companies. ICT is fully responsible for the lossless information transfer, not only on a regional level but also globally, therefore, it is an urgent call for these departments to realize that people communicate, perceive visual information and process data in different ways, depending on the cultural paradigm they have as their mental programming.

Acknowledgments The paper is a part of the SPEV 2020 Project run at the Faculty of Informatics and Management at the University of Hradec Kralove, Czech Republic, and the author would like to thank the student Ales Berger for his help with collecting the data of the research.

References

1. Alpaydin E (2016) Machine learning. The New AI. MIT Press
2. Buckland M (2017) Information and society. MIT Press, Cambridge
3. Clark A et al (2010) The handbook of computational linguistics and natural language processing. Chichester: Blackwell
4. Drigas A, Koukianakis L, Convergence of Culture and ICTs: E-Culture. https://www.researchgate.net/publication/255811400_Convergence_of_Culture_and_ICTs_E-Culture. Accessed 10 Oct 2019
5. Hofstede G (1984) Cultural dimensions in management and planning. Asia Pacific J Manag. Springer, 81–99
6. Hofstede G et al, Cultures and organisations. Software of the Mind. http://testrain.info/download/Software%20of%20mind.pdf. Accessed 10 Oct 2019
7. Ball, Donald A, Mcculloch, Wendel H (1996) International business. the challenge of global competition. New York: McGraw-Hill, Irwin
8. Bartlett, Christopher et al (2008) Transnational management. Text, Cases, and Readings in Cross-Border Management. New York: McGraw-Hill, Irwin
9. Deresky H (2008) International management. Managing Across Borders and Cultures. Text and Cases. New Jersey: Pearson Education
10. Pikhart M (2018) Multilingual and intercultural competence for ICT: Accessing and assessing electronic information in the global world (MISSI 2018). Advances in Intelligent Systems and Computing, 833, 2019, pp 273–278. 11th International Conference on Multimedia and Network Information Systems, MISSI 2018; Wroclaw; Poland; 12 September 2018 through 14 September 2018. ISSN 2194-5357. https://doi.org/10.1007/978-3-319-98678-4_28
11. Pikhart M (2020) Aspects of intercultural communication in it: convergence of communication and computing in the global world of interconnectedness. In LNEE 590:251–256
12. Pikhart M (2019) Multilingual and intercultural competence for ICT: accessing and assessing electronic information in the global world. Adv Intell Syst Comput 833:273–278
13. Pikhart M (2018) Technology enhanced learning experience in intercultural business communication course: a case study. In: Emerging technologies for education. third international symposium, SETE 2018, Held in Conjunction with ICWL 2018, Chiang Mai, Thailand, August 22–24, 2018, Revised Selected Papers. Book Series: Lecture Notes in Computer Science, Editors: Tianyong Hao, Wei Chen, Ha, Publisher: Springer International Publishing, Print ISBN: 978-3-030-03579-2, Electronic ISBN: 978-3-030-03580-8

14. Pikhart M (2019) Intercultural business communication courses in european universities as a way to enhance competitiveness. In: Proceedings of the 32nd international business information management association conference, pp 524–527
15. Klímová B, Pikhart M (2020) Cognitive and applied linguistics aspects of using social media: the impact of the use of Facebook on developing writing skills in learning english as a foreign language. Eur J Investig Health Psychol Educ 10(1):110–118
16. Pikhart M, Klímová B (2019) Utilization of linguistic aspects of Bloom's Taxonomy in blended learning. Educ Sci 9(3):235
17. Berger A, Klímová B (2019) Mobile application for the teaching of English. In: Park J, Loia V, Choo KK, Yi G (eds) Advanced multimedia and ubiquitous engineering. MUE 2018, FutureTech 2018. LNEE, 518, 1–6
18. Cheung SKS (2014) A case study on the students' attitude and acceptance of mobile learning CCIS 2014 (pp 45–54). Heidelberg: Springer
19. Klímová B, Berger A (2018) Evaluation of the use of mobile application in learning English vocabulary and phrases—A case study. In: Hao T, Chen W, Xie H, Nadee W, Lau R (eds) Emerging technologies for education. SETE 2018. LNCS, 11284, 3–11
20. Klimova B, Simonova I, Poulova P (2017) Blended learning in the university English courses: case study. In: Cheung S, Kwok L, Ma W, Lee LK, Yang H (eds) Blended learning. New Challenges and Innovative Practices. ICBL 2017. Lecture Notes in Computer Science, vol 10309. Springer, pp 53–64
21. Kukulska-Hulme A, Shield L (2008) An overview of mobile assisted language learning: from content delivery to supported collaboration and interaction. ReCALL 20(3):271–289
22. Mehdipour Y, Zerehkafi H (2013) Mobile learning for education: Benefits and challenges. Int J Comput Eng Res 3(6):93–101
23. Muhammed AA (2014) The impact of mobiles on language learning on the part of English foreign language (EFL) university students. Procedia Soc Behav Sci 136:104–108

Hierarchical Similarity Hash for Fast Malware Detection

Sunoh Choi

Abstract Hundreds of thousands of new malicious files are being created every day. Existing pattern-based vaccine engines cannot detect these new malicious files. To solve these problems, artificial intelligence based malicious file detection methods have been proposed. However, artificial intelligence based malicious file detection method has a disadvantage that takes long time because it requires dynamic analysis. We can use similarity hashes to solve these problems and find similar files. However, it also takes a long time to compare similarity hashes when there are many files. To solve this problem, this paper proposes a method to generate hierarchical similarity hash based on M-tree.

Keywords Similarity hash · M-tree

1 Introduction

Today, hundreds of thousands of new malicious files are created every day. It is difficult to detect them by existing pattern-based vaccine engines. To solve these problems, artificial intelligence based malicious file detection methods are being developed [1]. However, artificial intelligence based malicious file detection methods can take a long time because they may require dynamic analysis of files.

Similarity hash is proposed to solve this problem [2]. Similarity hashes differ from cryptographic hashes such as MD5. In the cryptographic hash, if a file is modified a bit, the hash value is completely different. However, similarity hashes in similarity files have similar hash values. So, given a variant file, we can quickly find a similar file using similarity hashes.

However, it can also take a long time to compare similarity hashes when the number of files is high. To solve this problem, this paper proposes a method to generate hierarchical similarity hash. Using hierarchical similarity hash, we can reduce the time to find similar files from O(n) to O(log n).

S. Choi (✉)
Computer Engineering Department, Honam University, Gwangju, South Korea
e-mail: suno@honam.ac.kr

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_19

Section 2 introduces a method to create a similarity hash and calculate the difference of the similarity hash value. In Sect. 3, we propose a method to create hierarchical similarity hash.

2 Similarity Hash

The purpose of the similarity hash is to find a variant of a file. Especially, in the case of malicious files, hundreds of thousands of variant malicious files are created every day, making it difficult to detect them with an existing pattern-based vaccine engine. Similarity hash can help to solve this problem. When there are existing files known to be benign or malicious, if a new file has similar similarity hash values, it can be judged to be similar to the existing file.

The property of the similar hash is as follows. A cryptographic hash such as MD5 has a completely different hash value, even if the file is slightly different. However, in the case of similarity hashes, similar files have similar hash values and others have different hash values.

A typical example of a similarity hash is Trendmicro locality sensitive hash (TLSH) [2]. How to create a TLSH is as follows. First, a byte string of the file is processed using a sliding window of size 5 to compute the bucket count array. Second, we calculate quartile points q1, q2, and q3. Third, we construct the digest header values. Fourth, we construct the digest body by processing the bucket array.

First, given a byte string of a file as shown in Fig. 1, it is counted in a 128-dimensional array using Pearson hash [3]. That is, the value computed by the Pearson hash through a windows of size 5 is matched to one of the 128 buckets and the count of the bucket is incremented by one.

Second, we calculate quartile points. 75% of bucket counts are greater than or equal to q1 and 50% of bucket counts are greater than or equal to q2 and 25% of bucket counts are greater than or equal to q3.

Third, as shown in Fig. 2, the first byte of the first 3 bytes of the similarity hash is the checksum, the second byte is the length of the file, and the third byte is calculated from q1, q2 and q3 as follows.

$$\begin{aligned} \text{q1\_ratio} &= (\text{q1} * 100/\text{q3}) \text{ MOD } 16 \\ \text{q2\_ratio} &= (\text{q2} * 100/\text{q3}) \text{ MOD } 16 \end{aligned}$$

Finally, if the value of each bucket is less than or equal to q1, put 00 into c_i, and if it is less than or equal to q2, put 01. If it is less than or equal to q3, put 10. Otherwise, put 11. TLSH is created through these four steps.

A method for calculating the similarity difference between two files using two similarity hashes is as follows. First, calculate the difference between the similarity hash header values. Second, calculate the difference between the similarity hash body values.

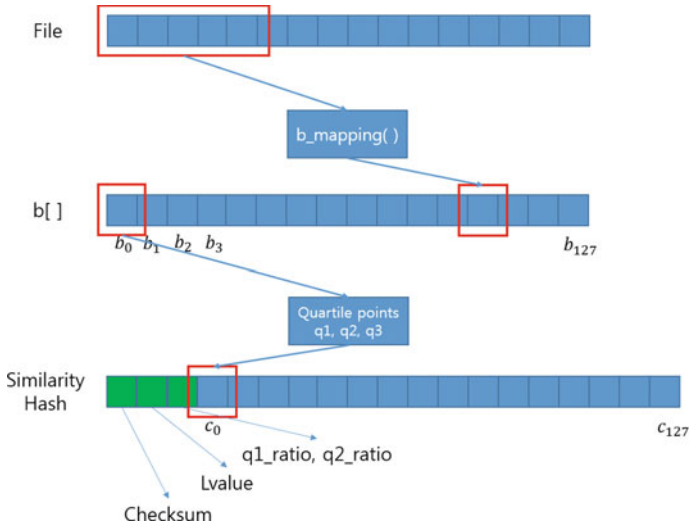


Fig. 1 How to create TLSH

The difference in the header values of the similarity hash is calculated as follows.

```
Function distance_headers(tX, tY)
  ldiff = mod_diff(tX.lvalue, tY.lvalue, 256)
  q1diff = mod_diff(tX.q1ratio, tY.q1ratio, 16)
  q2diff = mod_diff(tX.q2ratio, tY.q2ratio, 16)
  diff = ldiff * 12 + q1diff * 12 + q2diff * 12
  if (tX.checksum != tY.checksum) diff++
```

The difference in the body values of the similarity hash si calculated as follows.

```
Function distance_bodies(tX, tY)
  For i = 0 to 127 {
    d = abs(x[i] - y[i])
    if (d == 3) diff = diff + 6
    else diff = diff + d
  }
```

Finally, the difference between the two similarity hash values is calculated as the sum of the difference between the similarity hash header values and the difference between the similarity hash body values.

3 Hierarchical Similarity Hash

We can use similarity hash to find similar files to the new malware. However, when there are a lot of files, it takes so long time. In order to reduce the time, we propose hierarchical similarity hash (HSH) based on M-tree [4].

As shown in Fig. 2, the root node has two child nodes Oa and Ob. They are called routing objects. Each routing object Oi has its own covering r(Oi). It means that all its descendant nodes are within r(Oi). T(Oi) is the subtree whose root node is Oi.

Then, suppose that the query node is Q and we want to find all nodes which are within r(Q). M-tree has an important property. When d(Oi, Q) is greater than r(Q) + r(Oi), all nodes in T(Oi) can be pruned for search since all nodes in T(Oi) are not within r(Q) by the triangle inequality. By using HSH, we can find similar files to the new malware.

We can find similar files to the new malware by using HSH as follows.

(1) When the child node of the root node is Oi, check whether d(Oi, Q) is less than r(Q) + r(Oi).
(2) If yes, repeat the step 1 to check the child node Oj of the node Oi.
(3) If no, we do not need to check the child nodes Oj of the node Oi.

For example, suppose that the root node has two child nodes Oa and Ob in Fig. 2. Since d(Oa, Q) is greater than r(Q) + r(Qa), we do not need to check the child nodes of the node Oa. However, since d(Ob, Q) is less than r(Q) + r(Ob), we have to check the child nodes of Ob recursively.

HSH is generated as follows. When a new object On is inserted, we have to find a node Oi whose covering r(Oi) is greater than d(Oi, On). However, when there is no such a node, we have to increase the covering r(Oi) as small as possible.

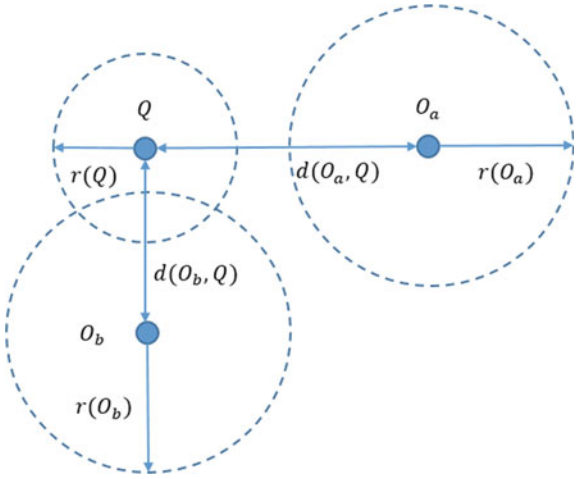


Fig. 2 Avoiding distance computations

In the future, we will implement HSH. Then, we will check how long the HSH will reduce the time to find similar files to the new malware compared to the case when we do not use HSH.

4 Conclusions

In this paper, we propose HSH for similarity hash. Similarity hashes are used to find similar files when new ones are given. However, finding a similar file when there are many files is a difficult problem. To solve this problem, we can use HSH. In the future work, we will show how long HSH will reduce the time to find similar files to the new malware.

Acknowledgments This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019R1G1A11100261).

References

1. Kaspersky (2017) Machine Learning for Malware Detection. http://media.kaspersky.com/en/enterprise-security/Kaspersky-Lab-Whitepaper-Machine-Learning.pdf
2. Oliver J, Cheng C, Chen Y (2013) TLSH—a locality sensitive hash, fourth cybercrime and trustworthy computing workshop
3. Pearson PK (1990) Fast hashing of variable-length text strings. Commun ACM
4. Ciaccia P, Patella M, Zezula P (1997) M-tree: an efficient access method for similarity search in metric spaces. In: VLDB

A Study on Dropout Techniques to Reduce Overfitting in Deep Neural Networks

Hyun-il Lim

Abstract Recently, neural networks are used in various applications to solve complex problems. The neural network creates models that can solve problems through training from data. However, overfitting problems in training models have a significant impact on the performance of a neural network. In this paper, the effect of dropout to reduce overfitting in neural networks is analyzed through experiments. It also analyzes the effect on the performance of neural networks according to the number of nodes in layers that is an important factor in designing neural network models. The results of these experiments will help to design the dropout rates and the number of nodes in layers to apply neural networks in various applications.

Keywords Neural network · Dropout · Overfitting · Machine learning

1 Introduction

In a recent computing environment, a variety of problems are being solved using machine learning approaches [1]. Machine learning is a way to find a solution to a specific problem by generating a model by learning a large number of previously known training data. To apply such problems to machine learning methods, it is important to design a learning model and to have sufficient training data. Machine learning generates a customized model to solve a problem in the process of training from the existing known data. So, the generated model can predict accurate results for new data as well.

One of the difficult problems to solve in machine learning is to reduce the overfitting that occurs during the training of data. Overfitting [2] of learning data means less accuracy when applied in real-world environments. This can occur primarily when training data contains a lot of noise, or the machine learning model is too complex and learns too much about the noise contained in training data. One way to reduce this problem of overfitting is to apply dropout [3–6] in the learning process

H. Lim (✉)
Department of Computer Engineering, Kyungnam University, Changwon, Gyeongsangnam-do 51767, South Korea
e-mail: hilim@kyungnam.ac.kr

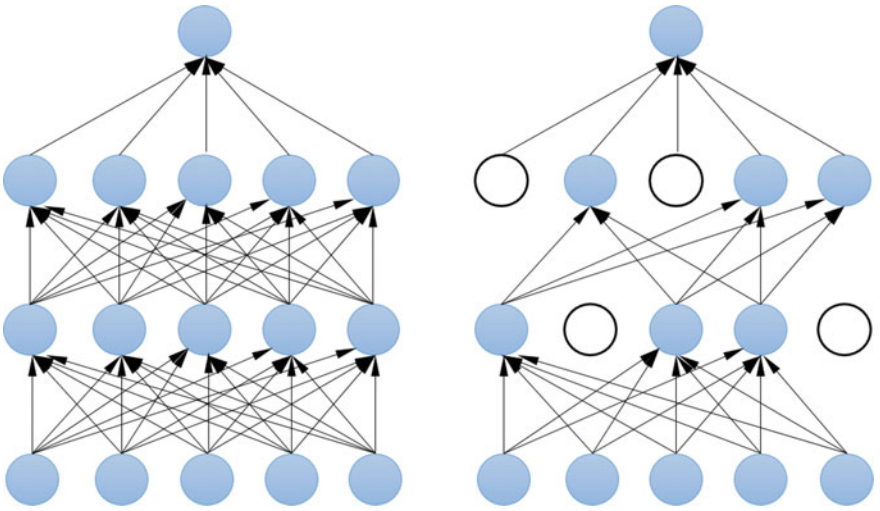
J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_20

of neural network models. In this paper, we analyze the effects of the neural network model according to the composition of the nodes of a deep neural network and the dropout rates, which make up an important factor in the design of a neural network. The results shown in this paper may be reflected in the design of neural networks and dropout to reduce overfitting in applying neural networks.

2 The Concept of Dropout in Deep Neural Networks

When a neural network model is trained with training data, the weight of each node is updated to make the model customized to the training data. The training might become too dependent on the training dataset. In such a case, the generated model will not give satisfactory results in prediction or classification problems. This is because the model is overfitted to the training data. To overcome this problem, the dropout technique can be applied in the phase of training a neural network model.

The idea of a dropout technique is to temporarily remove nodes from the original neural network based on probability in the phase of training the model. By applying dropout to neural network models, we can get less overfitted models for training data. In applying dropout to a neural network model, we assign a value to decide the probability of a node being excluded from the network in the training phase. So, the probability value is the ratio of excluded nodes among the total number of nodes in a layer of a neural network.



(a) A standard deep neural network. (b) After applying dropout.

Fig. 1 An example of applying dropout to a deep neural network

Figure 1 shows an example of applying dropout to hidden layers of a deep neural network. A standard deep neural network model is fully connected to the nodes in each layer of a neural network. In applying dropout to the neural network, each node is decided to be excluded from the network according to the dropout rate. After applying dropout to the neural network, the excluded nodes and the related edges are excluded from the network.

It is generally known that applying dropout to a neural network can reduce overfitting of the neural network. On the other hand, the technique can be more effectively applied in the design of a neural network model if we can understand how differently affect the accuracy and loss according to the dropout rates. So, it is also an important factor to decide the number of nodes in the design of a neural network.

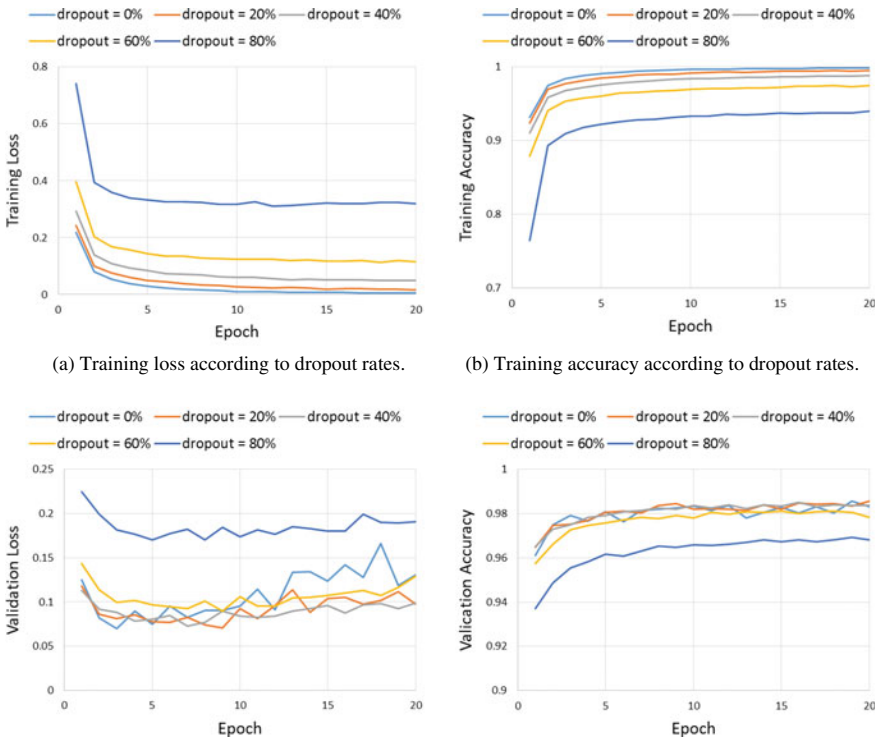
3 The Effect of Applying Dropout in a Deep Neural Network

In this section, we perform experiments to verify the effect on accuracy and loss according to the dropout rates and the effect of applying dropout according to the number of nodes in layers of a neural network.

3.1 The Effect of Dropout Rates

This experiment has been performed with a neural network model for recognizing the handwritten digits of the MNIST database [7]. The neural network model has two hidden layers, and the number of nodes in the hidden layer is designed to be 1024. As the training progresses, the accuracy and loss are measured when the dropout rates change from 0 to 80% for each hidden layer.

Figure 2 shows the results of experiments to verify the effect of applying dropout to a neural network according to dropout rates. The results show the effect on loss and accuracy as training of the model progresses. We can find that the training loss and accuracy is improved as the training epochs increase. The higher the dropout rate, the worse the training loss and accuracy. It is because higher dropout rates exclude the more nodes in training a model, and it disadvantageous in generating a neural network model customized for the training data. On the other hand, validation loss and accuracy show slightly different results. While the model without dropout technique shows the best result among all models, we can see that the validation loss increases rapidly after five epochs of the training phase. This means that the neural network model is overfitted to the training data at that time. Thus, the validation loss of this experiment shows the best results when the dropout rates are 20 or 40%. If the dropout rate is higher than the threshold value, we can confirm that the neural



(a) Training loss according to dropout rates.

(b) Training accuracy according to dropout rates.

(c) Validation loss according to dropout rates.

(d) Validation accuracy according to dropout rates.

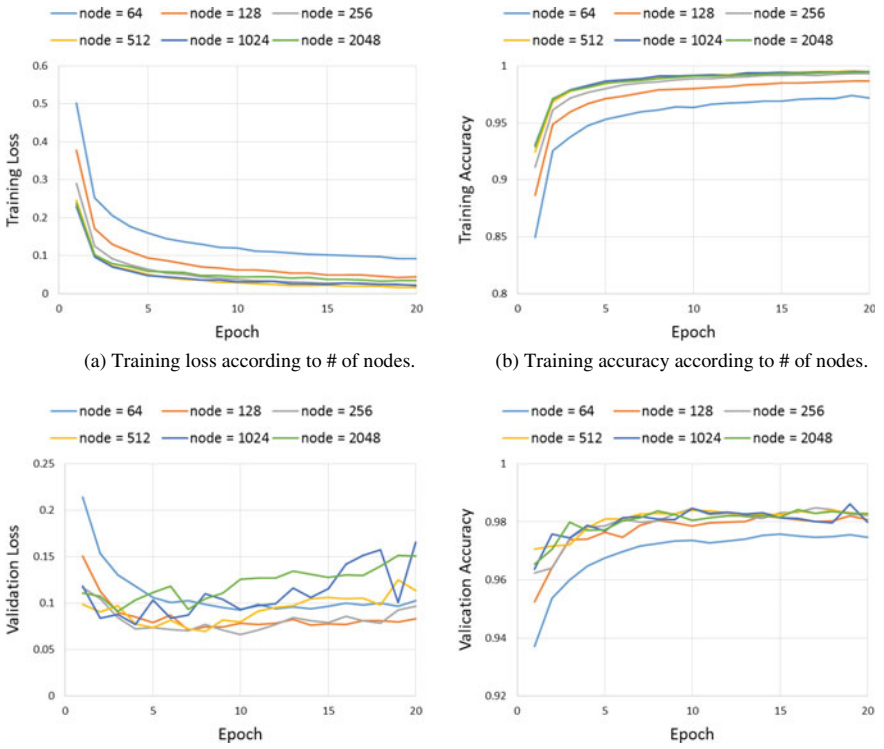
Fig. 2 The results of experiments for verifying the effect of applying dropout to a neural network according to dropout rates

network model cannot be satisfactorily trained as the number of excluded nodes increases by the dropout technique.

From the results of the experiment, training loss shows the best result when the dropout is not applied to the neural network because a model without dropout can reflect all nodes in training the model. However, as training progresses, validation loss increases because of overfitting. By applying dropout to the neural network, we can confirm that the validation loss reduces the overfitting in the model. However, excessive dropout rates can degrade the accuracy as many nodes are excluded in training the model.

3.2 The Effect of Dropout According to the Number of Nodes

In this section, we perform experiments to measure the effect of applying dropout to the neural network according to the number of nodes in layers of the network. The dropout rate is 20% and the number of nodes is increased from 64 to 2048 for



(a) Training loss according to # of nodes.

(b) Training accuracy according to # of nodes.

(c) Validation loss according to # of nodes.

(d) Validation accuracy according to # of nodes.

Fig. 3 The results of experiments on the effect of applying dropout to a neural network according to the number of nodes in layers

each hidden layer of the neural network. In this experiment, we evaluate whether a complex neural network with a large number of nodes is more advantageous for training. Figure 3 shows the experimental results for evaluating the effects of applying dropout according to the number of nodes in hidden layers. The training loss and accuracy are showing better results as the number of nodes increases, as seen in the experimental results of Sect. 3.1. This is because a complex model with a large number of nodes is advantageous in training the model that is customized for their training data. On the other hand, we can confirm that the validation loss increases after 5 epochs when the number of nodes is 2048. This means that complex models with a large number of nodes are overfitted to the training data even when dropout is applied to the model. When the number of nodes is 64, we can confirm that the validation accuracy does not increase by more than a certain level because the model does not have enough number of nodes for training. The results of this experiment show that the model has the most stable validation loss and accuracy when they have 128 or 256 nodes in their hidden layers.

From the results of the experiment, training loss of the model with the largest number of nodes shows the best result because it has the most complex structure. However, as training progresses, validation loss increases rapidly because of overfitting to the training data. On the contrary, reducing the number of nodes helps to reduce overfitting. However, if the number of nodes is not large enough to train the model, performance will no longer improve because the model cannot be trained sufficiently to training data.

3.3 Discussion

In designing a neural network model, we should consider several factors to improve the performance of the results. In these experiments, we show how the application of dropout is affected to reduce overfitting in the neural network model and how the performance is affected depending on the number of nodes in layers of neural networks. Therefore, to design a good neural network model, it is required to understand the performance impact of the dropout rates and the number of nodes and to reflect the effect in the design of a neural network. This is because complex models that simply increase the dropout rates and the number of nodes can rather degrade neural network performance.

4 Conclusion

It is important to reduce overfitting that adversely affects the performance of neural networks. In this paper, we verify that the performance of a neural network is dependent on the number of nodes and dropout rates. From the results of the experiment, it is necessary to effectively control overfitting because it increases validation loss. Therefore, it is important to determine an appropriate dropout rate and the number of nodes in designing neural network models to apply to real-world problems. The results of these experiments can be considered for applying dropout and determining the number of nodes in designing neural networks to reduce overfitting for various data analyses. They are also expected to be used for further studies to reduce overfitting in neural networks.

Acknowledgments This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Education) (No. NRF-2017R1D1A1B03034769).

References

1. Murphy KP (2012) Machine learning: a probabilistic perspective. The MIT Press
2. Salman S, Liu X (2019) Overfitting mechanism and avoidance in deep neural networks. arXiv: 1901.06566
3. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 15:1929–1958
4. Baldi P, Sadowski P (2013) Understanding dropout. In: Proceedings of the 26th international conference on neural information processing systems, pp 2814–2822
5. Duyck J, Lee MH, Lei E (2014) Modified dropout for training neural network
6. Labach A, Salehinejad H, Valaee S (2019) Survey of dropout methods for deep neural networks. arXiv:1904.13310
7. LeCun Y, Cortes C, Burges C (2020) The MNIST database of handwritten digits. http://yann.lecun.com/exdb/mnist/

Improvement Efficiency of Fingerprint Database Using Mathematical Methods

Jan Budina, Martin Zmítko, and Tomáš Kozel

Abstract Indoor localization based on iBeacons needs fingerprint database full of data and requires it to maintain it in good condition. During the database filling subjective anomalies indicating a certain dependence of fingerprint data, was found. These subjective findings need to be investigated, verified by mathematical statistics and understood why they arise. The following article is based on the case study and the data of the Faculty of Informatics and Management of the UHK. The article focuses on analyzing data using advanced statistical methods and determining whether the influence of data is affected by accuracy. The goal of this article is to find out if there is any dependency between individual fingerprints, use those results in location algorithms and verify those results in real environment, how this dependence affects further measurements. The discussion describes algorithms that improve database quality and verify it.

Keywords Fingerprint · Beacons · Indoor localization · Statistics · Database · Algorithm · Analysis

1 Introduction

The iBeacon Bluetooth system [1, 2] works with the fingerprint database for indoor positioning [3]. We can also consider using WiFi [4, 5], but it is not accurate as Bluetooth systems. Every mobile device that wants to determine a location detects it from both the iBeacon and by comparing it with existing fingerprints in the database [6]. After the position is scanned via application of web interface [7–9], another new

J. Budina (✉) · M. Zmítko · T. Kozel
Faculty of Informatics and Management, Department of Informatics and Quantitative Methods, The University of Hradec Kralove, Hradec Kralove, Czech Republic
e-mail: jan.budina@uhk.cz

M. Zmítko
e-mail: martin.zmitko@uhk.cz

T. Kozel
e-mail: tomas.kozel@uhk.cz

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_21

fingerprint is stored in the database [10, 11]. In addition to location information, the fingerprint also includes complete information about the manufacturer and type of phone that is necessary for further analysis. During testing of this system, some deviations occurred that occurred accidentally. Mostly 4 types of mobile devices were used for measurement. If the data in the database was limited to only one type with which the measurement itself took place, the results were much more accurate and mainly without any deviations. If data from all devices were used, there were deviations there. This resulted in the suspicion of some dependence between the scanned data and the type of scanner. Analyzing the acquired data helps to improve the system. The improvement itself may be in the form of editing the acquired data (separating some samples) or designing a mechanism for automatic repairs of fingerprints directly in the database [12, 13]. The subject of the analysis will be the interest of the mobile device and its hardware on the fingerprints [5, 14]. In this case, we can identify a number of variables that we can take into account, such as the measured attenuation, the number of scans, the measurement time, and more. This analysis will contribute to the further continuation of specific research at the Faculty of Informatics, which aims at refining the indoor localization using Bluetooth beacons, in this case, iBeacons [12]. The main methods for comparing these data will be ANCOVA and MANOVA statistical methods [15, 16]. The MANOVA method is a multivariate analysis of variance [17, 1]. Its use is appropriate just when we have two or more dependent variables. This method also includes significance tests that apply to individual dependent variables.

Article should answer the basic questions:

- Do the independent variables have a impact on the dependent variable?
- What is the relationship between the dependent variables?
- What is the relationship between the independent variables?

The ANCOVA [18] method is a covariance analysis that combines variance analysis (ANOVA) and regression analysis (RA). Both of these methods are in the form of a general linear model, with both assuming that the dependent variable (Y) is a random contiguous variable and differs in the assumptions about the properties of the independent variable (X). The basic idea of ANCOVA [7] is to extend the basic model of scattering analysis to a model that contains quantitative contiguous variables that affect the explained variables.

2 Materials and Methods

The aim of these methods is to purify the dependence of the explanatory variable on the chosen factors from the various influential influences (covariates) and to prove the dependency of the explained variable on the factors. All tests are performed in IBM SPSS Statistics [1, 2]. In the following chapters, these tests will be described in more detail and the results evaluated. Fingerprint collecting took place on the FIM building, where data from the database was collected using the application in the

phone. The specific investigation of possible combinations, i.e. the relation of an independent and dependent variable that may interest us, is defined as follows:

- Mobile phone tag: average scanning time BLE, average scan time Wi-Fi
- Mobile phone tag: average BLE attenuation, average Wi-Fi attenuation
- Mobile phone tag: number of unique BLE/Wi-Fi scans unique
- Mobile phone tag: influence of BLE on the quality of Wi-Fi scanning on different phone brands.

2.1 Initial Experiment Setup

The basic experiment starts on Faculty of Informatics and Management at the University of Hradec Kralove, where were iBeacon transmitters placed. Optimal position [3] of iBeacons was made by an article [4]. Experiment tries to get best position information in the building using iBeacons and fingerprint database [5]. Unfortunately, the accuracy of fingerprints is not constant, but it depends on the device, which was used for measuring. This finding was important to make another research and testing with our existing fingerprint database [6] contain thousands of rows of data. Some mobile phone brands have better sensitivity than others. It was necessary to distinguish each brand and analyze data between and compare them.

Initial setup was brands: Sony, Vodafone, Samsung, Google, Bluetooth transmitters: 30, fingerprint records: 4067, measuring time: 6000 ms.

MANOVA Hypothesis

A, Attenuation scans BLE/WIFI versus Brand.

At the 5% significance level ($\alpha = 0.05$), we will test the zero H0 hypothesis ie that there is no difference between the measured BLE/Wi-Fi attenuation by different cell phone brands compared to the alternative H1 hypothesis that there is a difference.

H0: $= \mu$google $= \mu$samsung $= \mu$sony $= \mu$vodafone.
H1: $= \mu$google $<>$ samsung or μgoogle $<> \mu$vodafone or μsamsung $<> \mu$vodafone or μsamsung $<> \mu$sony or μsony $<> \mu$vodafone or μgoogle $<> \mu$sony.

And similar hypotheses are B, C.

B, Number of unique scans BLE/WIFI versus Brand.

C, Average scan time scans BLE/WIFI versus Brand.

3 Results

A, MANOVA for BLE/WIFI Attenuation Versus Brand

The box diagram from Fig. 1: box diagram in the case of BLE points to the rejection of the H_0 hypothesis because one median from the selection does not fall into the other's selections. This is, however, only an indicator. The actual investigation will take place on the basis of further tests. In the multivariate test we see Sig = 0.00 for $p < 0.0005$. Therefore, it can be concluded that at a significance level of 5%, a statistically significant difference in measured values (BLE, Wi-Fi) is found between at least two mobile phone tags. In the next step, it will be necessary to examine which variables are involved in this difference, whether all or just some.

We have now done our own MANOVA calculation. This test tells us if any of the dependent variables are different. Wi-Fi is $F(3.880) = 23$; $p < 0.0005$; $\eta^2 = 0.073$ and for BLE is $F(3.880) = 39.5$; $p < 0.0005$; $n2 = 0.119$. It means it's statistically significant. $\chi^2_{0.95}(3.880) = 2.62$ i.e. $39.5 > 2.62$, and thus, at the level of significance $\alpha = 0.05$, we reject the zero hypothesis about the insignificant influence of the individual variables. We can conclude that there is a different signal attenuation on BLE and Wi-Fi, according to the brand. The resulting conclusion can be determined by multiple comparisons. Our goal is to find out which variables there is a significant difference. In addition, differences in the level of individual variables between groups will have to be revealed. It is true that sig < 0.0005 here is a statistically significant difference. MANOVA shows that there is a statistically significant difference between the average Wi-Fi attenuation between Google and Sony (sig 0.00), but there is no difference between Google and Samsung (sig., 0.383) or Google and Vodafone). Furthermore, there is no statistically significant difference between the Samsung brand compared to other brands. However, there is a statistically significant difference between Sony and Google and Vodafone mobile phones and their type of chipsets. In this respect, it is important from this point of view that Sony

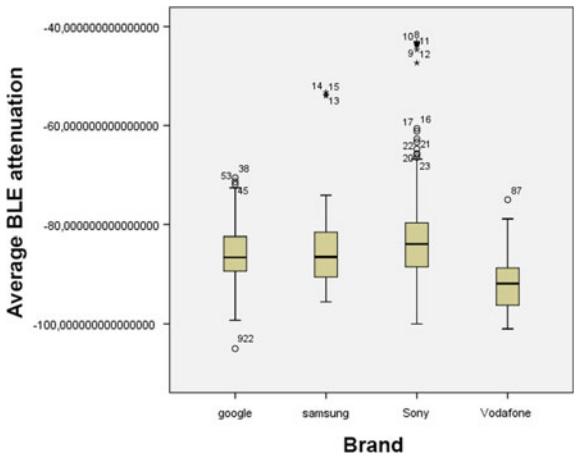


Fig. 1 Box diagram

has a statistically significant difference over other Wi-Fi signals. For this reason, this mobile device will be further monitored. In the case of measuring BLE attenuation, it is interesting that the Vodafone brand differs significantly from other brands, i.e. there is a statistically significant difference between Vodafone and other brands when measuring the BLE attenuation. From this point of view, it can be concluded that for a Vodafone mobile brand, it will be necessary to further strengthen the iBeacon device's signal, possibly taking into account other parameters in the system. Test of between subjects effects determines whether any of the dependent variables vary (whether there is a statistically significant difference). Wi-Fi, $F(3.918) = 148{,}486$; $p < 0.0005$; $n2 = 0.010$ and for BLE is $F(3.918) = 4.156$; $p < 0.0005$; $\eta^2 = 0.022$. Which means it's statistically significant. $\chi^2_{0.95}(3.918) = 2.61$ i.e. $148{,}486 > 2.62$ and $4.156 > 2.61$, and thus, at the level of significance $\alpha = 0.05$, we reject the zero hypothesis about the insignificant influence of the variables. We can conclude that there are differences between the BLE and Wi-Fi scanning times.

Based on these findings, an algorithm was created that uses previous findings to improve the fingerprint database. This algorithm consists of two parts, the analyzer and the optimizer. The analyzer will run at regular intervals over the fingerprint database and will try to analyze it and test its status. The obtained knowledge is passed on to the output as input for the second part of the so-called optimizer, which will ensure the actual optimization of the database or its part. By separating the analyzer, we ensure that the overall mechanism has minimal impact on the overall database load. The analyzer outputs are further processed by the optimizer, which is the second component of the system. The optimizer also works on the fingerprint database. In this case, there is no recurrence during the day. This component is scheduled to run only once a day to improve database health. All algorithms such as extreme separation, clustering are linked to each other and share their output data. The optimizer can be schematically described in Fig. 2.

The verification and testing of these algorithms took place over an existing data table, which was constantly filled with new data for a week. First part was filling database with regular data without any cleaning and optimizing database without any interruption. Second part was about filling database with data, which are optimized

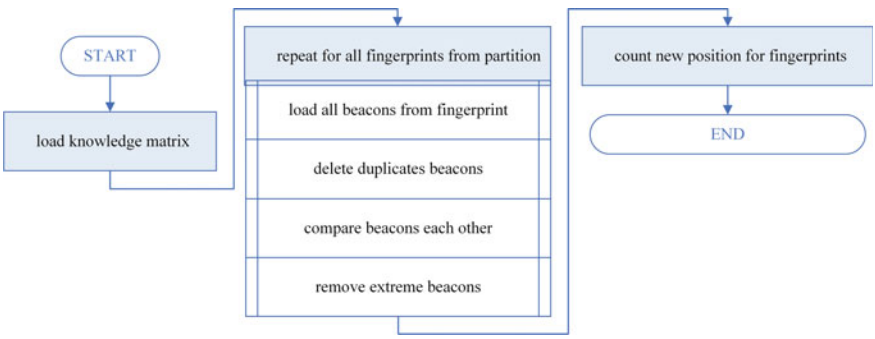


Fig. 2 Optimalizer schema. Source: author

with previous knowledge about phone vendors. It should be visible some difference between clean data and optimized data in database. Measuring was at the same place for both database filling. Extreme separation and clustering algorithms were chosen for the experiment. Testing was performed on a basic sample of 50 fingerprints and every day 50 new ones were added. Each evening, the above-mentioned database optimization algorithms were run, and then the improvement was measured using a test agent. An overall improvement of 12% with the help of database optimization is a very good result. It can be assumed that running the experiment for more than one week would improve again, but the percentage increase in accuracy would decrease until it stabilized at a static value. This phenomenon was further observed outside the experiment, where improvement stopped at 13% after one month. All data were compared with the same measurement without algorithms, that consider different phone vendors.

4 Discussion

It is clear from the statistical findings above that the data measured in the fingerprint database are influenced by the device that measured them. This degenerates the quality of the database to a certain extent because every device, even if it may have a chip from the same manufacturer, affects the quality in one direction or the other. In order to avoid these problems in the future, it is necessary to store the type and manufacturer of each device. Based on this information, the database can be better optimized. Given the current situation, where there are thousands of types of phones, data cannot be flattened according to their types. In order to process the data better and better, it is necessary to use certain methods that have self-learning algorithms. Nowadays, it is mostly neural networks [9–11] that are able to correctly set their parameters and adapt the input data so that it is as constant as possible and is not influenced by external influences, based on existing data and presented model set.

The result of this finding is therefore the implementation of self-learning algorithms that will continually browse and check the database and clear it from inaccuracies caused by the use of different phone brands.

5 Conclusion

The aim of this work was to investigate the relationship between individual brands of mobile phones with the measured dependent variable based on the data obtained from the database. Testing has shown that there are indeed statistically significant differences, especially when comparing the Wi-Fi method [12, 13] to BLE. In this case, the BLE method appears to be more prone to individual phone parameters than Wi-Fi. If we take a closer look at BLE technology in relation to mobile phones, the previous analysis shows that, for example, with regard to measured attenuation, the

Vodafone brand is very prone to measurement and shows significantly worse performance. This fact requires work in the future and, if necessary, proposing adjustments to prevent this problem if it turns out that the system is for any reason whatsoever.

The overall statistical test summary also confirms that there is a dependency between phone manufacturers and Bluetooth and WiFi devices. This is partly due to the principle that phones are designed when multiple manufacturers use the same hardware for commonly used components. The result continues to affect the scanning of a Bluetooth device that broadcasts at the same frequency.

Results of the experiment help to increase indoor localization reliability. In our case, we need to add weight to every record by the phone brand. The more accurate records should have a higher weight coefficient than less accurate records. These weights will use our indoor localization system to deliver position data by phone brand and weight index in the database [14], which theoretically increase position accuracy.

Acknowledgments This work was supported by the SPEV project, financed from the Faculty of Informatics and Management, University of Hradec Kralove.

References

1. Laerd statistics: one-way MANOVA in SPSS statistics. Laerd statistics, 18 May 2016. https://statistics.laerd.com/spss-tutorials/one-way-manova-using-spss-statistics.php
2. Field AP (2013) Discovering statistics using IBM SPSS statistics, ISBN: 978-1-4462-7458-3 978-1-4462-7457-6
3. Kriz P, Maly F, Kozel T (2016) Improving indoor localization using bluetooth low energy beacons. Mob Inf Syst s. 1–11. https://doi.org/10.1155/2016/2083094
4. Budina J, Klapka O, Kozel T, Zmitko M (2015) Method of iBeacon optimal distribution for indoor localization. In: Christiansen H, Stojanovic I, Papadopoulos GA (eds) Modeling and using context, vol 9405. Springer International Publishing, Cham, s 105–117
5. Zhang C (2013) Institute of Electrical and Electronics Engineers, a IEEE systems, man, and cybernetics society (ed) 2013 IEEE 14th international conference on information reuse and integration (IRI 2013): San Francisco, California, USA, 14–16 August 2013 [including workshops]. IEEE, Piscataway, NJ
6. Sacca D, Wiederhold G (1985) Database partitioning in a cluster of processors. ACM Trans Datab Syst roč. 10, č. 1, s. 29–56, bře. https://doi.org/10.1145/3148.3161
7. Hebák P (2007) Vícerozměrné statistické metody. Informatorium, Praha
8. Simple Linear Regression. Eberly Colleage of Science, 25 May 2016. https://onlinecourses.science.psu.edu/stat501/node/250
9. A. Baldominos, E. Albacete, Y. Saez, Isasi P (2014) A scalable machine learning online service for big data real-time analysis. In: 2014 IEEE symposium on computational intelligence in big data (CIBD), Orlando, FL, USA, December 2014, s 1–8. https://doi.org/10.1109/CIBD.2014.7011537
10. Kriesel R, Kriesel D (2015) Neural networks. https://page.mi.fu-berlin.de/rojas/neural/chapter/K8.pdf
11. Yegnanarayana B (2006) Artificial neural networks. Prentice-Hall of India, New Delhi
12. Pei L et al (2012) The evaluation of WiFi positioning in a Bluetooth and WiFi coexistence environment. In: 2012 ubiquitous positioning, indoor navigation, and location based service

(UPINLBS), Helsinki, Finland, October 2012, s 1–6. https://doi.org/10.1109/UPINLBS.2012.6409768
13. Baqer M (2010) Enabling collaboration and coordination of wireless sensor networks via social networks. In: 2010 6th IEEE International conference on distributed computing in sensor systems workshops (DCOSSW), Santa Barbara, CA, June 2010, s 1–2. https://doi.org/10.1109/DCOSSW.2010.5593272
14. Budina J, Zmitko M (2018) Idea of learning context system using iBeacons. In: 2018 ELEKTRO, Mikulov, May 2018, s 1–6. https://doi.org/10.1109/ELEKTRO.2018.8398378
15. Keselman HJ, Huberty CJ, Lix LM,Olejnik S, Cribbie RA, Donahue B, Kowalchuk RK, Lowman LL, Petoskey MD, Keselman JC, Levin JR (2016) Statistical practices of educational researchers: an analysis of their ANOVA, MANOVA, and ANCOVA analyses. Rev Edu Res 68(3):350–386
16. O'Brien RG, Kaiser MK (1985) MANOVA method for analyzing repeated measures designs: an extensive primer. Psychological Bulletin 97(2):316–333
17. Drápela K (2013) Vícerozměrná analýza rozptylu (MANOVA)—její předpoklady a využití'
18. Miller GA, Chapman JP (2001) Misunderstanding analysis of covariance. J Abnorm Psychol 110 (1):40-48

Academic Paper Personalized Search Based on Interest Subject Tracking and Search Performance Evaluation

Sangjun Han, Sukhyung Lee, Namgyu Kang, Heeseok Choi, and Sungpil Choi

Abstract Web services that provide academic papers present information based on their own search algorithms. However, since it only judges the relevance between the query and the document without understanding the user's interest field and the retrieval intent, there is a limit to presenting the optimal information for each individual. In this paper, the user search intention at the time of retrieval is identified based on the subject classification and it is determined as the area of interest. In particular, we propose personalized search as a way to provide information suitable for individual interests when homonym or different meanings are used in different research fields. The proposed personalized search performance based on the subject classification tracking showed higher search performance than the comparison target.

Keywords Search personalization · Search performance evaluation · Interest field tracking · Academic paper

S. Han (✉) · S. Lee · N. Kang · H. Choi
Convergence Service Center Korea, Institute of Science and Technology Information, 245 Daehak-ro, Yuseong-gu, Daejeon 34141, South Korea
e-mail: sjhan@kisti.re.kr

S. Lee
e-mail: skyi@kisti.re.kr

N. Kang
e-mail: ngkang@kisti.re.kr

H. Choi
e-mail: choihs@kisti.re.kr

S. Choi
Department of Library and Information Science, Kyonggi University, 154-42 Gwanggyosan-ro, Yeongtong-gu, Suwon 16224, South Korea
e-mail: spchoi@kgu.ac.kr

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_22

1 Introduction

NDSL (web service), which provides academic papers, is a National Science and Technology Information Center established and operated by the Korea Institute of Science and Technology Information [1]. It provides about 100 million procedures and journals. However, providing the same search results in the same form and ranking for the same query language is difficult to satisfy the user's demand with different interests. Therefore, in this research, user interest field which is changing every moment is tracked in real time based on the DDC(Dewey Decimal Classification), and personalization of thesis search result is made by using it. Until recently, many researches on retrieval result reorientation and information recommendation using user profile analysis, query expansion, collaborative filtering, and similar document clustering techniques have been performed [2–7]. However, there is a lack of research on personalized search considering the changing situation of individual search intention. The purpose of this study is to improve user's satisfaction in terms of information utilization by personalizing the information of the user by tracking the user's changing areas of interest. In order to verify the accuracy of personalized search results, we compared the search performance with NDSL and proved the superiority of this study.

2 Personal Interest Subject Tracking

The questionnaire was used to derive the degree of interest in the academic paper of each function utilization pattern to be used for tracking user interest. Table 1 shows the degree of interest in academic paper by user behavior. Classification information

Table 1 Interest score in academic paper information by usage activities

Paper information usage pattern	Interest score
View full paper (PDF view)	81
Full paper copy application	80
Search by exact title	74
Bookmarks	71
Purchase (Payment)	69
Print detail page	68
Go to detailed view	67
Add to cart	63
Export	32
SNS Sharing	46
View the similar documents	35
Keyword search (many search results)	23

to be used as a personal interest field uses the code name of DDC. This table shows the results of response to the questionnaire surveyed on the degree of interest in the paper by the user behavior. The interest score was quantified by the weighted average calculated by factoring the degree of interest and the number of users responding.

Each time the defined 12 usage activities occur, the interest score of the DDC code corresponding to the user interest field is accumulated and updated. Over time, a variety of behaviors will occur in a complex manner, and the number and scores of subdivisions corresponding to the area of interest will be increased and updated. The score update for each area of interest is calculated on a daily basis, and the relative ratios are calculated using the scores of all the classified categories determined in the area of interest as shown in Eq. (1).

$$Ratio\,of\,interest\,t_p = \frac{DAYSCORE_p}{\sum_{i=1}^{s} DAYSCORE_p} \times 100,\ where\,s\,is\,the\,number\,of\,SEQ \quad (1)$$

In the personalized search, top-3 DDC codes with high relative proportion among individual interest fields are utilized. In the search results, a technique of re-adjusting the rank value of academic papers having the corresponding code is utilized. Figure 1 shows the result of tracking individual interests after performing various activities in an online service, such as searching some papers or viewing full paper. The user's interest in this case is ultimately the subject of life sciences, artificial intelligence, and clinical medicine. In the figure, the x-axis represents the number of user's paper-use actions, and the y-axis represents the ratio of user interests.

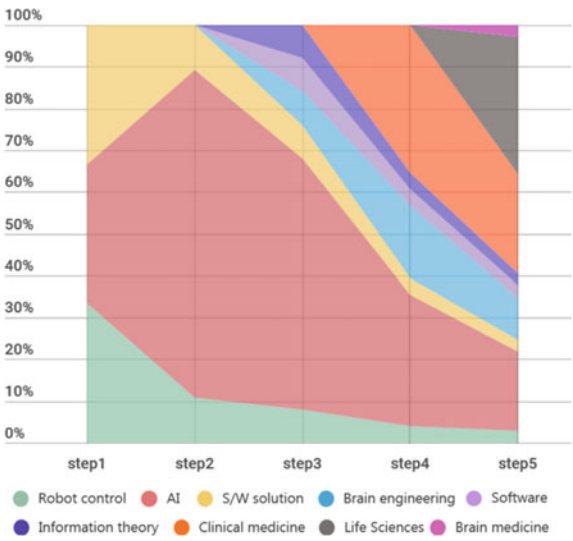


Fig. 1 Personal interest tracking result example

3 Academic Paper Personalized Search by Interest Tracking

Individual interests of interest tracked by DDC code are reflected in the ranking of thesis search results. When a search engine receives a query and a personal interest, the search engine generates a result set suitable for the query. Then, rearrangement is performed by placing papers having specific interest information among the search result sets at the top. However, based on the statistics that the search rate of the information in the top three of the first page is more than 50% and the search effort of the desired information is hardly performed at the third page or more in the search portal, We applied a method of assigning weights to 30 papers and re-ranking them to be included in the ranking of thirty of the total search results. 'Cell' means the basic structure and activity unit of an organism in biology. In mobile communication engineering, it refers to a region covered by a single communication base station. Also, in the area of semiconductors and software engineering, 'cell' refers to the smallest unit that stores data in memory. The results of the paper search in the NDSL using the query 'cell' are the results that do not take into consideration individual interests and are shown in Table 2. Table 3 shows the search results of the user whose

Table 2 Search results for 'cell' in NDSL(not personalized)

Seq	Title of paper	Subject
1	Cell mediated immunity to allografts of fresh and treated bone	Occurrence/Neurobiology
2	Transfert graisseux enrichi en cellules souches: concepts actuels	Radiation technology
3	Cell-mediated immune responses to a killed Salmonella enteritidis vaccine: lymphocyte proliferation, T-cell changes and interleukin-6 (IL-6), IL-1, IL-2, and IFN-γ production	Occurrence/Neurobiology
4	Aspects actuels de la drepanocytose chez l'enfant au Gabon	Other Life Sciences
5	Associations jonctions-organites cellulaires dans le foie nEo-natal de souris	Semiconductor device/system
6	I-Cell disease (mucolipidosis type II) serum hydrolases in obligate heterozygotes	Molecular Cell Biology
7	I-Cell disease, mucolipidosis II: Pathological, histochemical, ultrastructural and biochemical observations in four cases	Occurrence/Neurobiology
8	Cell size and cell number in the wing of drosophilia melanogaster as related to parental ageing	Drug/drug development
9	Cell-wall structure in the growth-zone of Phycomyces sporangiophores—I. Model experiments and microscopical observations	Occurrence/Neurobiology
10	Stem cell therapy in veterinary dermatology	Occurrence/Neurobiology

Table 3 Search results for 'cell' of users who interested in biology by proposed method

Seq	Title of paper	Subject
1	Dihydroergotamine Tartrate Induces Lung Cancer Cell Death through Apoptosis and Mitophagy	Molecular Cell Biology
2	Hypoxia-induced expression of cellular prion protein improves the therapeutic potential of mesenchymal stem cells	Molecular Cell Biology
3	Promising Therapeutic Strategies for Mesenchymal Stem Cell-Based Cardiovascular Regeneration: From Cell Priming to Tissue Engineering	Stem cell biology
4	The prognostic value of interim and end-of-treatment PET/CT in patients with newly diagnosed peripheral T-cell lymphoma	Molecular Cell Biology
5	Bioinspired tuning of glycol chitosan for 3D cell culture	Stem cell biology
6	An 8-gene signature for prediction of prognosis and chemoresponse in non-small cell lung cancer	Molecular Cell Biology
7	MAPK Cascades in Guard Cell Signal Transduction	System Biology
8	A Genetic Screen Reveals Novel Targets to Render Pseudomonas aeruginosa Sensitive to Lysozyme and Cell Wall-Targeting Antibiotics	Microbiology/Parasitology
9	Biphasic activation of extracellular signal-regulated kinase (ERK) 1/2 in epidermal growth factor (EGF)-stimulated SW480 colorectal cancer cells	System Biology
10	Stepwise inhibition of T cell recruitment at post-capillary venules by orally active desulfated heparins in inflammatory arthritis	Molecular Cell Biology

interest is biology through the proposed personalized search method, and Table 4 shows the search results of the user whose interest is mobile communication, U-computing, and software. The papers presented by the proposed method reflect the personal interest field at the time of using the information.

4 Personalized Search Performance Evaluation

In order to evaluate the proposed personalized search performance, 5 researchers in computer and information science fields, 5 researchers in biotechnology and bio/medical field were selected and instructed to use personalized search service for about 2 months. Then, the retrieval performance evaluation results between the NDSL and the proposed personalized search were collected using the presented query words. First, Precision@10 was calculated by checking the correct answer document among the top 10 results of the papers search for 10 questions by experts. The correct answer document, which is meant here, is judged by whether there is relevance between the research area of the expert, the query word and the paper. In

Table 4 Search results for 'cell' of users who interested in mobile communication, U-computing, and software by proposed method

Seq	Title of paper	Subject
1	A Cell Searching Technique without Double Counting for a Mobile Station with Multiple Antenna Arrays in Millimeter Wave Cellular Communication Systems	Ubiquitous -computing platform/application technology
2	An Enhanced Clustering Method Based on Grid-shaking	Software
3	Stability of SiNX/SiNX double stack antireflection coating for single crystalline silicon solar cells	Software
4	Development of 2-d unstructured hybrid grid generation program using java applet	Software
5	Classification of Tumor cells in Phase-contrast Microscopy Image using Fourier Descriptor	Software
6	A Statistical Inter-Cell Interference Model for Uplink Cellular OFDMA Networks Under Log-Normal Shadowing and Rayleigh Fading	Ubiquitous -computing platform/application technology
7	Virtual Cell Beamforming in Cooperative Networks	Software
8	NDRG2 is one of novel intrinsic factors for regulation of IL-10 production in human myeloid cell	Ubiquitous -computing platform/application technology
9	Library Support in an Actor-Based Parallel Programming Platform	Software
10	SCAMP5 Links Endoplasmic Reticulum Stress to the Accumulation of Expanded Polyglutamine Protein Aggregates via Endocytosis Inhibition	Software

the next step, the search performance was evaluated by calculating the MAP (Mean Average Precision), which means the average value of the average document accuracy for each query word, because there are 10 query words used in the evaluation. The MAP indicates how many correct documents are ranked at the top among the top ten result documents. For reference, the MRR (Mean Reciprocal Rank) is not used as a measure of the accuracy of how much related documents are included in search results, but rather is used to determine how quickly and topically a single correct answer document is presented to the user. In this study, the improvement of search performance is aimed at presenting the most suitable document at the same time as presenting the document suitable for the individual interest field and the query at the top of the search result. Therefore, the accuracy of the search result was measured by assuming that the document that the user judged to be suitable for the field of interest and the query is the correct answer. We measured Precision@10 to determine how

Table 5 Precision@10 evaluation results and improvement rate

Query	NDSL	Proposed method	Improvement rate	Average
Virus	0.284	0.534	0.250	0.302
Robot	0.354	0.659	0.305	
Cell	0.298	0.704	0.406	
Window	0.366	0.656	0.290	
Bigdata	0.425	0.748	0.323	
Dementia	0.377	0.654	0.277	
Deep running	0.421	0.802	0.381	
Analysis	0.346	0.585	0.239	
Collapse	0.252	0.544	0.292	
Browsing	0.339	0.592	0.253	

many correct answer documents are included in the top 10, and measured MAP to determine how many correct document has top position in the top 10 rankings.

4.1 Precision@10 Evaluation and Analysis

Precision@K measures the number of correct answers among the top K search results. The ten evaluators performed a search on each of the ten query words, and judged that the document having relevance between the user's interest field, the query term, and the search result document at the time of the search was 'T(ture)'. In NDSL, only 3–4 of the top 10 result documents were judged to be correct, and in the proposed personalized search, 7 correct answers were identified among the top 10 result documents. Table 5 shows the result and the performance of Precision@10 is improved by 0.302 (87.28%) compared to NDSL.

4.2 MAP Evaluation and Analysis

We measured the MAP to determine how many correct documents in the Precision@K evaluation are located at the top area of the 10 search results. The MAP finds all of the Precision@n in rank position n (n = 1 to 10) checked in the correct answer document among the top 10 search results by ten query words, calculates the average, and then calculates the average thereof. Table 6 shows the MAP evaluation results and improvement rates. As a result, the proposed personalized search against NDSL improved by 0.307 (60.43%).

Table 6 MAP evaluation results and improvement rate

Query	NDSL	Proposed method	Improvement rate	Average
Virus	0.325	0.654	0.329	0.307
Robot	0.510	0.781	0.271	
Cell	0.580	0.712	0.132	
Window	0.382	0.824	0.442	
Bigdata	0.443	0.852	0.409	
Dementia	0.502	0.936	0.434	
Deep running	0.577	0.787	0.210	
Analysis	0.698	0.897	0.199	
Collapse	0.512	0.852	0.340	
Browsing	0.546	0.845	0.299	

5 Conclusion

There is an increasing demand for users who want to find and use appropriate information as much as the amount of information that increases. In providing the academic paper information, it is necessary to identify the user's intention or interest in the search, and present the appropriate information to the easy-to-use form or search ranking, rather than simply providing the information with just the fit between the query word and the document. In this study, personalized search based on the DDC is proposed to provide articles that meet the researcher's tendency or interests and have high relevance for the users. In addition to future research, it is necessary to further study the research of interest field tracking and personalization search techniques that can be applied to various fields of science and technology information, and to demonstrate its performance.

Acknowledgments This research was supported by the Development of S&T Knowledge Infrastructure Convergence Service Program of the Korea Institute of Science and Technology Information (KISTI) (K-20-L01-C07-S01).

References

1. National Digital Science Library (NDSL), http://www.ndsl.kr
2. White, Ryen W et al (2013) Enhancing personalized search by mining and modeling task behavior. In: Proceedings of the 22nd international conference on World Wide Web, pp 1411–1420
3. Wang W, Peery C, Marian, Nguyen TD (2012) Efficient multidimensional fuzzy search for personal information management systems. IEEE Trans Knowl Data Eng 24(9):1584–1597
4. Ali SH, El Desouky AI, Saleh AI (2016)A new profile learning model for recommendation system based on machine learning technique. Indonesian J Electr Eng Inform 4(1):81–92

5. Lei L et al (2014) Modeling and broadening temporal user interest in personalized news recommendation. Expert Syst Appl 41(7):3168–3177
6. Priya TS, Revathy P, Pradeesh T, Robin CRR (2012) Design and development of an ontology based personal web search engine. Procedia Technol 6:299–306
7. Kim Y-H, Kim H-G (2006) Ranking decision method of retrieved documents using user profile from searching engine. J Korea Inst Maritime Inform Commun Sci 10(9):1590–1595

Implementation of an Authentication System GTPass

Monsoon Khadka and Gi-Chul Yang

Abstract This paper describes on implementation of an efficient personal authentication scheme called 'GTPass' (Graphic & Text based Password system). GTPass establishes a bulwark against shoulder-surfing attacks and is easy to use. GTPass introduces a novel method of integrating numerical and graphical password authentication techniques. The important feature of GTPass is the use of graphics and numbers together at the same time for a system. GTPass is easy to use and robust against shoulder-surfing attacks.

Keywords Security · Authentication · Password

1 Introduction

So far, a variety of personal authentication systems have been introduced in order to provide easy, convenient and secure access to certain system. But the main problem we're facing today is that it's hard to remember complex passwords. Even if we manage to remember them, there is always a risk of exposure to others, which may result in unauthorized access to the system by unauthorized users. In a typical security system, passwords are in text format or pattern, which are slow to type and can be easily stolen or be guessed, once seen, by other. In such scenario we usually create a new password. If newly created password is short and easy to remember than there is high risk of exposure, so to avoid it we may create longer and more complex password. As a result, it may be difficult for the user to remember and may forget it.

Graphical passwords have been developed to solve the problem described above. The graphical password was first proposed by Blonder [1], and these days there are diverse schemes for graphical passwords [2–8] to replace text-based passwords. However, most of the graphical password systems are weaker against shoulder-surfing attacks than text-based password systems, since graphics are easier to remember than text.

M. Khadka · G.-C. Yang (✉)
Mokpo National University, Mokpo, Korea
e-mail: gcyang@mokpo.ac.kr

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_23

GTPass introduces a scheme that can lead to the development of the ultimate graphical password system [9]. GTPass increases usability while maintaining strong security. It increases usability by facilitating quick logins: a user only needs to type in numbers like a text-based password system. GTPass maintains strong security by using a correspondence relationship between graphics and numbers within the system.

2 GTPass System

This section describes workflow and structure of GTPass along with the GTPass system's interface. The general password space for GTPass was calculated by the formula

$$P = (X * Y)^N \text{ in } [9],$$

where X is the number of columns and Y is the number of rows and N is the number of password images selected. For example, consider a grid of seven columns and seven rows with a password of eight images. The password space is

$$P = (7 * 7)^8 = 33, 232, 930, 569, 601.$$

Clearly, GTPass can generate a very large password space. Thus, it can provide a strong defense against brute force attacks [9].

3 Workflow and Structure of GTPass

GTPass system works as follows. If a user selects 'register' in the interface than the system will direct the user into the Register Activity where the user is asked to click on choose Image which will display the images from the database in a random pattern. Then the user has to select those images which he/she want to keep as password (note: images on the screen will change their position randomly on every selection user makes). When the user selects any images than the system will store the string associated with that image into the data store and add them together. After which the system will arrange all the images in a random pattern and display it into the screen. This process keeps on repeating until the user presses the Done Button. After 'Done' is selected or pressed, the system will store the string/data stored in Data Store into Database. During the process of registering a password, if the user makes any mistake or selects the wrong image than he/she can just simply press the reset button and recreate the password again.

If a user selects Login than the system will direct the user into Login Activity where the user is asked to select the password images in the same way that they were selected while creating the password. While doing so, Images are displayed onto the screen in a random pattern. As soon as the user inputs coordinate of the password image and press next, the system will select the image on that position and store its string value into the Data Store and add them up. Then images are displayed onto the screen in a random pattern. This process keeps on repeating until the user selects Done Button. After Done Button is clicked, the system will retrieve the original password from the database and compared it to the string in the Data Store. If the string matches with the password, then the system will allow access for the user else "wrong password" message will be displayed and the user will be asked to start from the beginning. During this process, if the user makes any mistake or selects the wrong image than he/she can just simply press the reset button and re-enter the password again (Fig. 1).

3.1 Interface of GTPass

GTPass system is implemented with Java for Windows and Android (Fig. 2).

GTPass System can be described in two phases. They are:

1. **Creating password**: As In Picture [1], user has to select the picture that he/she want to add in their password. After every selection, position of pictures in the screen is shuffled at random. In case user mistakenly selects wrong picture or want to recreate password form the first then he/she can just press refresh button and recreate the password and have to enter backup password which can be used in case he/she forgets the password
2. **Login**: As in Picture [2], while entering password to get access to certain system, user has to enter the position of the picture from password as in screen by pressing button twice where first number will describe the position of the picture in column and second button will describe its position in row. After selecting two numbers, user has to press right arrow button which will inform the system that position of the picture is entered. After which, every picture in the screen will be shuffled randomly and the user has to follow the same process again and again according to the sequence of picture in password. In case miss-selection user can re-enter the position of current picture by pressing cancel button on the top right or re-enter the password by pressing the cancel button twice. After entering the entire password the user has to done (correct) button to get access to the system. In case, user forgets his/her password than can simply use back-up password to create new password.

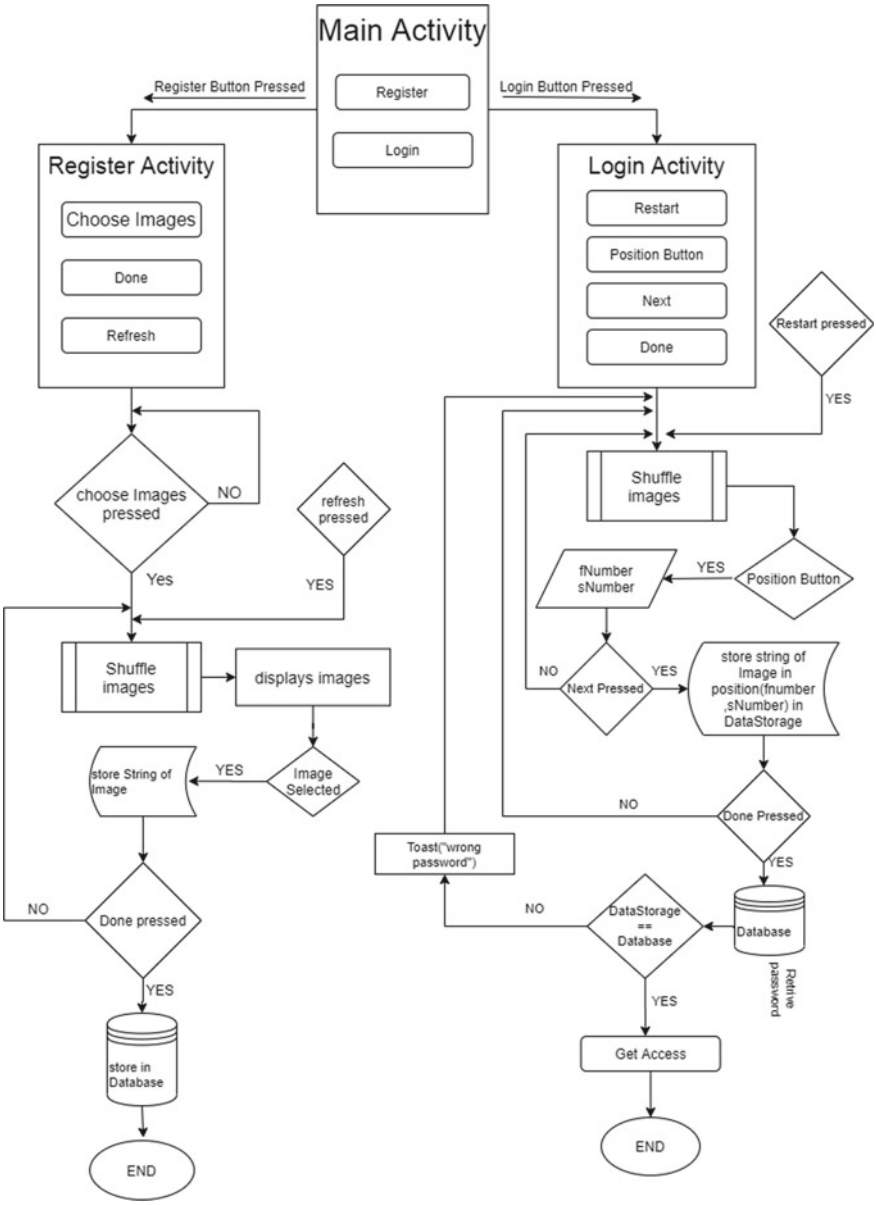


Fig. 1 Workflow and structure of GTPass

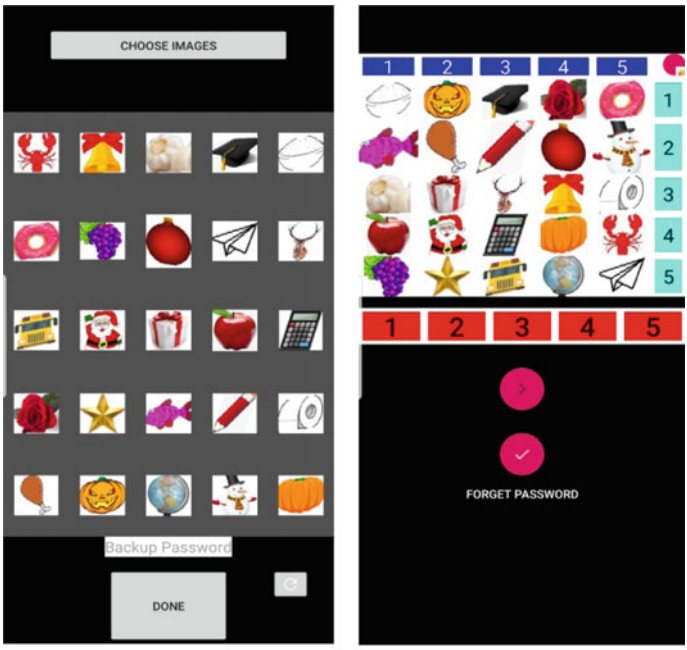


Fig. 1 Interface of GTPass

4 Conclusion

Password systems should be secure and easy to use. However, text-based systems require passwords that can be difficult to remember, while graphical systems require passwords with lengthy input times. In addition, shoulder-surfing attacks are a difficulty common to both.

This paper described implementation details of a multi-modal password system called GTPass. GTPass is a secure personal authentication system while it is easy to use. GTPass uses numbers to enter a password instead actual password entity itself (i.e., graphic image). Hence, it is cumbersome to steal someone's password through shoulder-surfing attack. It provides quick login time with strong security. Also, the GTPass system is highly secure and resistant to various attacks, since it does not require input password entity itself to login. GTPass uses both graphics and text to capitalize on the advantages and eliminate the drawbacks of both systems. Accordingly, GTPass will be easy to use and highly resistant to a variety of attacks.

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

References

1. Blonder G (1996) Graphical passwords. Lucent Technologies, Inc., Murray Hill, NJ, U. S. Patent, Ed. United States
2. Dhamija R, Perrig A (2000) Deja Vu: a user study using images for authentication. In: Proceedings of 9th USENIX Security Symposium
3. Akula S, Devisetty V (2004) Image based registration and authentication system. In: Proceedings of midwest instruction and computing symposium
4. Weinshall D, Kirkpatrick S (2004) Passwords You'll Never Forget, but Can't Recall. In: Proceedings of conference on human factors in computing systems (CHI). Vienna, Austria: ACM, pp 1399–1402
5. Jansen W (2004) Authenticating mobile device users through image selection. In: Data security
6. Wiedenbeck S, Waters J, Birget JC, Brodskiy A, Memon N (2005) PassPoints: design and longitudinal evaluation of a graphical password system. Int J Human Comput Stud 63:102–127
7. Yang G (2019) Development status and prospects of graphical password authentication system in Korea. KSII Trans Int Inform Syst 13(11):5755–5772. https://doi.org/10.3837/tiis.2019.11.026
8. Gokhale A, Waghmare V (2015) Graphical password authentication techniques: a review. Int J Sci Res 4(7):279–285
9. Yang G (2020) A multimodal password system based on graphics and text. Eng Lett 28(2):300–305

Designing a Predictive Model for Efficient Learning Outcomes in Korean Education

Ji-Hoon Seo and Kil-Hong Joo

Abstract The present study extracted meaningful data from domestic students' accumulated past learning outcome data utilizing the Educational Data Mining analysis technique and processed the data to design a university entrance probability prediction model and derive indicators that can be utilized by learners in the prediction of university entrance examination results. The design model presented in the present paper proposed analysis modeling to semantically classify individual learners' training data tables based on the basic policy of Korean education for the entrance examination in order to create position models suitable for the attributes of prediction and predict university entrance examination routes.

Keywords Big data · Predictive model · Education data mining

1 Introduction

Currently, as the data related to diverse fields accords the whole range such as the fields of industry, culture, economy, and arts have become big data globally the big data analysis technique intended to extract meaningful information from massive data is becoming a big issue throughout the world [1]. In particular, since the big data analysis technique enables extracting meaningful data based on the structured data and unstructured data formed from massive data collection devices and utilizing the extracted data in prediction analysis in diverse fields, it enables efficient and realistic derivation of information. Therefore, gradual studies are conducted in the field of education in South Korea based on educational data mining analysis models for enhancement of learner centered education programs, and based on the studies, future

J.-H. Seo
Institute of General Education, Incheon National University, 119 Academy-ro, Yeonsu-gu 406-772, Incheon-si, Republic of Korea
e-mail: jihoon@inu.ac.kr

K.-H. Joo (✉)
Department of Computer Education, Gyeongin National University of Education, 155 Sammak-ro, Manan-gu, Anyang-si, Republic of Korea
e-mail: khjoo@ginue.ac.kr

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_24

prediction systems for learner support have been developed. As a result, the scope of studies is being reduced to find customized learning systems that fit the propensities of individual learners and accordingly, constraint conditions are formed in diverse scopes of studies in the field of education, which is problematic [2]. In the case of South Korea, currently, since the importance of software education in preparation for creative education and the fourth industrial revolution has been recognized, new curriculum subjects are newly organized and the educational system is rapidly changing. However, despite the effort of the government as such, the stream of education in South Korea has been generally focused on university entrance until now. Therefore, in the present paper, considering the domestic education system, indicators that can be utilized in the prediction of the results of domestic university entrance examinations are proposed by applying the data mining technique to extract meaningful data from learners' accumulated grade data. In step 1 of the present study, information on learners' grades and desired universities is classified and stored in the collection and storage place. In step 2, learners' gender and learning data are constructed. In step 3, a model to classify students' grades based on past data into groups of a total of four levels, which are High group, Normal group, Low group, and Failure group, is presented and the preprocessing process is carried out. In step 4, whether the examinee can enter the university desired by him/her is determined utilizing the groups by students' grade classified in step 3. The indicators for Educational Data Mining based prediction of success or failure in the entrance examination presented in the present paper were proved to be usable as indicators more effective for the prediction of the results of domestic learners' university entrance examinations because the reliability and performance of ID3 and C4.5, which are two data mining techniques used in the present study, were compared and analyzed and the results indicated that the prediction accuracy was excellent.

2 Related Works

2.1 Decision Tree Algorithm

Decision trees are used to classify data in data mining, and have an advantage that they can be directly used in decision making because the results can be easily understood since they are expressed in tree structures [3]. ID3, which is a representative algorithm of decision trees, is an alternative for NP-complete problems and is a method to classify data through entropy and information gain values and create trees utilizing the Greedy Algorithm based on the purity of the classified certain nodes. Currently, decision trees are known to be an element that becomes the basis of CHAID, CART, C4.5, etc., mainly used in data mining classification algorithms [4].

2.2 Korean Based Education Data Mining

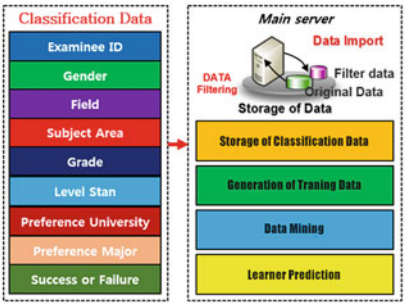
In South Korea, learner analyses are conducted in diverse areas for the purpose of cultivating creative talents. According to the results of such analyses, customized education programs suitable for individual learners are developed, and coding education and the free-semester system are introduced. Since problems in finding indicators of objective educational attainment and success or failure from the results of university students' liberal arts examinations in the prediction modeling currently utilized in the field of educational data mining were presented, a technique to compare and analyze decision tree algorithms such as C4.5, CART, and CHAID, which are classification techniques in data mining to explore determinants, was derived as a technique to solve the problems. However, this technique shows limitations to be used as a measure to judge learners' future oriented decision making. In addition, despite that the paradigm of domestic education is currently shifting toward advanced creative education, College Scholastic Ability Test centered rating scales are still necessary even now.

3 Design of Directivity Index for University Entrance

The technique proposed in the present paper extracts learning or training samples, which are the past learning data of the examinees wishing to enter the university. Training data were formed using the extracted data based on the classification model, and using the Attribute Selection Method and Cluster Analysis applicable to prediction branch points, the list of analysis of the possibility for examinees to enter the university and department prediction and selection model design were derived. The prediction accuracy was analyzed based on prediction selection modeling and a technique to judge success or failure in the entrance examination for the university preferred by the learner is presented to learners.

3.1 Classification of Learners Past Data

In step 1 of the indicator technique proposed in the present study, learning or training samples are extracted and classification models are created and imported into the database. In step 2, training data are created based on the classification model. By applying the model to the training data, attributes can be selected according to prediction branch points through the analysis procedure. Step 3 is a prediction step to judge success or failure. In the prediction step, indicators that would enable identifying success or failure based on learners' preferred university in the entrance examination are derived to judge where the learner can enter the university. With learners' classified information, the IDs of the elective subject areas corresponding



	1	2	3	4
	Score	gender	Level stan	university
1	0,90000	male	HIGH	S-UNI
2	0,90000	male	HIGH	I-UNI
3	0,70000	male	NORMAL	H-UNI
4	0,80000	female	HIGH	D-UNI
5	0,80000	female	HIGH	J-UNI
6	0,60000	female	NORMAL	H-UNI
7	0,50000	female	NORMAL	K-UNI
8	0,50000	female	NORMAL	C-UIN
9	0,40000	male	LOW	D-UNI
10	0,45000	male	LOW	J-UNI

Fig. 1 Procedure of the classification model and Test input data setting

to the College Scholastic Ability Test, learner identification IDs, and classification models according to genders were designed and in the case of learners' genders, rules were generated so that male learners would not be included in the prediction of entrance into women's universities. The indicator prediction technique proposed in the present study analyzes the high school grades of the examinee and the accuracy of the prediction of success or failure in the entrances examination for the university desired by the examinee to judge whether the examinee can enter the desired university. In the present study, data tables for 10 items according to the analysis procedure were designed. The tables are largely classified into learner information, university grade information, learner grade information according to prediction results (Fig. 1).

3.2 Creation of Rules for Learner Success or Failure Prediction Analysis

In the present study, for data mining processing for judgment of success or failure prediction, the measures by step as follows were presented based on the IF-THEN Rules (Fig. 2).

First, the past learner data and university data that will be analyzed are inputted. Second, learners' genders are classified. Since there are women's universities where only women can enter and universities where both male and female students can enter, students are first divided by gender in order to present a measure to exclude male learners from grading in the data mining stage.

Third, learners are classified according to their high school grades. Students are classified into a total of four groups according to their levels, that is, students with grades not lower than 0.8 into the high group, those with grades not lower than 0.5 but lower than 0.8 into the normal group, those with grades not lower than 0.3 but lower than 0.5 the low group, and those with grades lower than 0.3 into failure group that cannot enter university. In the case of grades as such, rules were created based on the rolling admission grades of domestic universities and data on groups by university were also stored in the same way. Step 4 is the final step to determine

Field - Rule			
Humanity	Natural	Industrial	Second Foreign L
A-01	A-01	A-01	B-ALL
A-02	A-02	A-02	C-ALL
A-03	A-03	A-03	D-ALL
B-01	C-01	D-01	E-01
B-02	C-02	D-02	E-02
B-03	C-03	D-03	E-03
B-04	C-04	D-04	E-04
B-05	NULL	D-05	E-05
B-06	NULL	NULL	E-06
B-07	NULL	NULL	E-07
B-08	NULL	NULL	E-08
B-09	NULL	NULL	E-09
B-10	NULL	NULL	NULL

Grade Code	Minimum(%)	Maximum(%)
1st	0	4
2nd	4	11
3rd	11	23
4th	23	40
5th	40	60
6th	60	77
7th	77	89
8th	89	96
9th	96	100

Fig. 2 Grade code standard data by area of learners and Learner grade standard by area

Table 1 Basic rule of entrance examination success or failure prediction indicators

*IF-THEN RULE
STEP 1:"exam_test" set s //input data setting
The set s of 25 examples with PASS and if then
STEP 2 Gender rule
IF GENDER = F→all, ELSE GENDER = M → Wuni.Fail
STEP 3 Score rule [Max Score 1]
IF Total Grade Score (Score >=0.8) → High
ELSE IF Total Grade Score (Score >=0.5 <=0.8) → Normal
ELSE IF Total Grade Score (Score >=0.3 <=0.5) → Low
ELSE Total Grade Score (Score <0.3) DOWN → Failure
STEP 4 UNIVERSITY Collection no. Success of failure rule
IF High→{S, H, D, Y, K, I, G, N, C, J, G, N, B} = Y
IF Normal→{I, G, N, C, J, G, N, B} = Y
ELSE IF {S, H, D, Y, K} = N, IF Low→{J, G, N, B} = Y
ELSE IF {S, H, D, Y, K, I, G, N, C} =N
ELSE Failure→{S, H, D, Y, K, I, G, N, C, J, G, N, B}=N

whether the examinee can enter his/her first preference university using the data classified through grouping by student grade in step 3. Individual universities were identified with alphabet capital letters. The domestic university entrance examination success or failure prediction indicators in the present study become to predict whether students can enter university or not through a four-step process. The IF-THEN rules proposed in the present study are as shown in Table 1.

4 Performance Evaluation

The present experiment utilized the data of 1,000 unspecified learners, and arbitrary values for the past grade data were set and constructed in the database. To derive diverse analysis results, scores by area according to subjects were joined to learner information based on the rules of the College Scholastic Ability Test in 2015 to derive

student ID	Score	gender	Level stan	university	success of failure
s001	0.9	male	HIGH	S-UNI	Y
s002	0.9	male	HIGH	I-UNI	Y
s003	0.7	male	NORMAL	H-UNI	N
s004	0.8	female	HIGH	D-UNI	Y
s005	0.8	female	HIGH	J-UNI	Y
s006	0.6	female	NORMAL	H-UNI	N
s007	0.5	female	NORMAL	K-UNI	N
s008	0.5	female	NORMAL	C-UIN	Y
s009	0.4	male	LOW	D-UNI	N
s010	0.45	male	LOW	J-UNI	Y

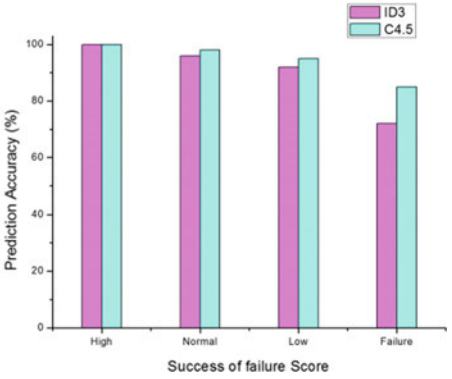


Fig. 3 Prediction analysis result and comparison of prediction accuracy

prediction results. University groups were classified utilizing the rolling admission information of a total of 60 universities by collecting sample data from the top 20 universities in the country, 20 middle ranked universities and 20 low ranked universities, including five universities which are women's universities, industrial universities, and the Korea Military Academy. Some of the university entrance examination prediction analysis results derived utilizing the prediction indicators presented in the present study are as shown in (left side of the Fig. 3). The performance evaluation in the present study was conducted using C4.5 and ID3. Prediction accuracy was analyzed based on learnrs' grade indexes for university groups, which are High, Normal, Low, and Failure. The results of comparison of prediction accuracy are as shown in (right side of the Fig. 3). In general, in ID3, the accuracy of prediction of Failure was shown to be low but high prediction accuracy not lower than 90% was derived for other items.

The prediction accuracy of ID3 and that of C4.5 were compared and analyzed and according to the results, prediction accuracy of C4.5 was 4% higher than that of ID3 because the prediction accuracy of ID3 was shown to be 90.5% on average while that of C4.5 was shown to be 94.5% on average.

5 Conclusion

The present study presented a Directional Index for University decision-making analysis technique that utilizes education data mining to generate a classification model based on the learning data of the examinee and determines the success or failure in the university entrance examination of the learner through attribute selection. As a result of verification by comparison and analysis of ID3 and C4.5, it was found that the average accuracy of C4.5 was 4.5% higher than that of ID3. Therefore, it is necessary to carry out follow-up studies focusing on the precise classification of data according to the entrance examination system and the performance evaluation using

diverse data mining techniques. Hereafter, we plan to conduct studies to present the diversity of classification of the attributes of data through previous studies presented in the present paper, and to improve the accuracy of learning data.

References

1. Wu X, Zhu X, Wu G-Q, Ding W (2014) Data mining with big data. IEEE Trans Knowl Data Eng 26(1):97–107
2. Russom P (2011) Big data analytics. IBM TDWI research. Fourth quarter
3. Idri A, Elyassami S (2011) Applying fuzzy ID3 decision tree for software effort estimation. Int J Comput Sci Issues (IJCSI) 8(4), no 1:131–138
4. Bhardwaj R, Vatta S (2013) Implementation of ID3 algorithm. Int J Adv Res Comput Sci Soft Eng (IJARCSSE) 3(6):845–851

Development of Korean Tourist Information Platform Based on Big Data

Ji-Hoon Seo and Kil-Hong Joo

Abstract The Korean Wave contents industry, which is growing steadily around the world, is now attracting foreign tourists interested in K-POP, movies, dramas, and cosmetics mainly in Asian countries. However, when foreign tourists have actually entered South Korea, they cannot fully enjoy the tourist courses planned by them before they go back to their countries in most cases due to time and spatial constraints. In addition, the return rate of foreign tourists who visited South Korea once was only 38% as of last year, which is much lower compared to 61% of the neighboring country Japan. Therefore, in this study, public data for individual regions opened by the government were collected by fully utilizing the rich IT infrastructures, which are an advantage of South Korea, to construct an integrated big data based tourism information visualization platform based on the web interlocked with tourism information service systems including GIS based maps thereby implementing a system to conduct opinion mining analysis with tourism destination evaluation comments collected through the web and accumulate or deduct the points of relevant tourism destinations according to positive or negative reputations.

Keywords Big data · Korean wave contents · Integrated tourism information platform · Opinion mining

J.-H. Seo
Institute of General Education, Incheon National University, 119 Academy-ro, Yeonsu-gu 406-772, Incheon-si, Republic of Korea
e-mail: jihoon@inu.ac.kr

K.-H. Joo (✉)
Department of Computer Education, Gyeongin National University of Education, 155 Sammak-ro, Manan-gu, Anyang-si, Republic of Korea
e-mail: khjoo@ginue.ac.kr

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_25

1 Introduction

Currently, in South Korea, along with economic growth, the size of the tourism industry is growing thanks to diverse contents including the Korean Wave. In fact, according to national statistics, the size of the domestic tourism industry as of 2015 was 73 trillion won, which amounted to 2.5% of the GDP. The results as such can bring about the synergy that can enhance the national brand and value because foreign currency incomes improve the balance of international payments and activate the economy leading to the recovery of the domestic economy, additional job creation, and employment effects. According to a survey by the Korea Tourism Knowledge Information System, the ratio of the tourism industry in the domestic employment market in 2015 was estimated to be 6.2% in the entire employment market. Therefore, the added value of the tourism industry cannot be ignored. However, the growth trend of the domestic tourism industry has slowed down recently due to the aftermath of the deterioration of the political and diplomatic relations with the neighboring country, but differently from the foregoing, the Korean Wave content industry has continued to grow year by year, thanks to Korean Wave contents such as K-POP, movies, dramas, and cosmetics. This indicates that Korean Wave content has grown beyond a simple field of the content industry to become an axis with the potential to add power to the national economic growth engine. The Korean Wave contents as such are intensively grafted on and utilized in business marketing. Therefore, for continuous qualitative growth of the domestic tourism industry, 1:1 customized interactive information service platforms from the position of tourists who are consumers should be urgently developed instead of the unilateral 1:N service platforms. Therefore, in this study, taking notice of the necessity of qualitative spread of tourism information services, a converged web app tourism information guide service platform on which spatial information data and big data analysis technologies have been grafted that can provide only necessary tourist information to tourists who are users was developed. This platform is a big data based integrated tourism information visualization platform that can visualize domestic tourist attractions and surrounding tourist infrastructure facilities, including Korean Wave contents to provide active customized information to individual tourists, collect, store, and classify user comment data on the web or applications to build a sentiment dictionary thereby conducting opinion mining analysis of unstructured data of clients to obtain accurate positive and negative reputations of tourism information, and visualize the foregoing to provide diverse and useful information on domestic tourism destinations or tourism infrastructures.

2 Related Work

2.1 Unstructured Data Analysis Technology

The core of the tourism information platform pursued in this study is not simply distinguishing between good and bad tourist attractions like social networking services, but clearly discerning which infrastructures were excellent and what were bad with a more fundamental perspective [1, 2]. Therefore, in this study, an unstructured data analysis technology is necessary to construct a platform based on factual grounds by collecting the reviews of clients who visited tourism destinations and analyzing the unstructured data [3, 4]. As a related example, text mining techniques enable obtaining resultant values more than those obtained through information searchers by extracting significant information data, identifying linkage with other information data, and finding the categories of texts [5].

2.2 Tourist Information Platform

In the case of domestic tourist information platform, there is a platform called every corner of South Korea developed by the Korea Tourism Organization. This platform complies with web accessibility and is a platform where foreigners of various nationalities can look into information on domestic tourist attractions. Although location information and photo information are presented, there is no analysis system that enables identifying individuals' objective opinions and reputations. In the case of the neighboring country Japan, the Tourism Agency provides travel information by collecting big data from diverse information such as messages posted on social media sites by about 700,000 tourists in eight regions of the country.

3 Proposed Method

3.1 Data Collection and Storage Construction

In this study, open data on tourist information provided by the government were collected. The collected data were information on top 100 domestic tourist attractions, Korean Wave drama and movie locations, and restaurants around tourist attractions (Fig. 1).

The collection servers of these data were separately classified and the GPS coordinate values of individual location data were constructed into Pivot tables as a requirement for registration of the location data of tourist attractions in individual regions on the open street map. This study utilizes location-based services to provide

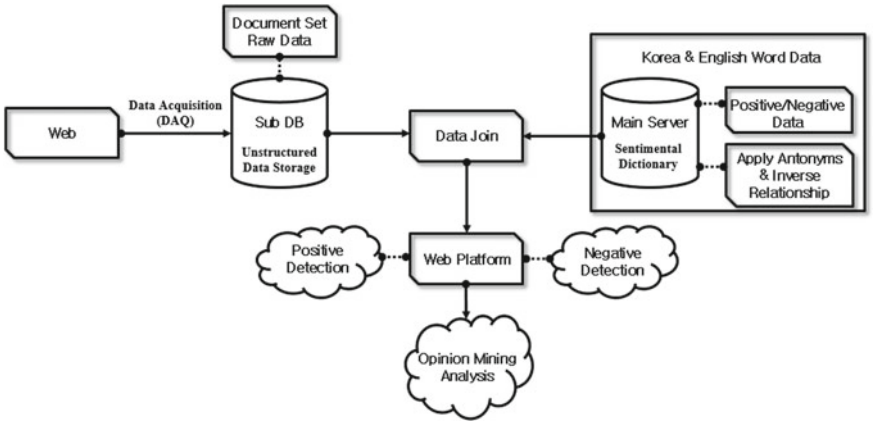


Fig. 1 Opinion mining data collection and DB modeling process

information on Korean Wave content such as search services for regions in 100 m to 1 km from the relevant tourist attraction, movie and drama locations, K-POP concert dates, and specialty stores for Korean wave products, and information on amenities for tourists such as nearby lodgings and restaurants. In addition, as shown in Fig. 2, centering on the current location of the user indicated as a green point on the mobile device screen, an open street map was implemented so that the results of opinion mining analysis can be easily understood by anybody by visualizing the degree of positive reputation into blue points and the degree of negative reputation intro red points in relation to the analysis of the reputation of the relevant tourist attraction.

The platform developed in this study enables users to leave a review of about 150 characters for each of the tourist attractions visited by them. All the review data as

Fig. 2 Basic concept of user participatory tourist attraction reputation information and positive & negative reputation data point visualization screen on GIS MAP

Table 1 Sentiment word filtering rules of the sentiment dictionary for extraction of sentiment word candidate groups

Sentiment word filtering rules for construction of the sentiment dictionary
1. Remove special characters, English words, and stop words
2. Remove meaningless terns and one-character texts
3. Distinguish between homonyms/synonyms
4. In the case of abbreviations and coined words, only the terms listed in the Wikipedia and the Korean dictionary are reflected

such are transmitted to the main server that manages websites, where the data are collected and stored, and the data are continuously accumulated and managed as big data.

3.2 Construction of Sentiment Dictionaries for Opinion Mining Analysis

For opinion mining analysis in this platform, those words that were extracted through preprocessing of meaningful words are classified into sentiment word candidate groups. In this study, in order to improve the accuracy and reliability of the resultant values from opinion mining analysis, the top 20% of the words selected as the sentiment word candidate groups were extracted and finally selected as the sentiment words in compliance with a series of word filtering rules as shown in Table 1. The data distributed in the subgroups consist of sets of words with high weights or meaningful words, but most of them are classified into neutral words or words close to extremity when they are tagged as positive, negative, neutral, or other words and extremity analysis is conducted.

4 Performance Evaluation and Analysis

The figure below is the main screen of the Korean tourist information platform. Korean and English can be switched, and switching to mobile form is possible because the platform was constructed into HTML5 based responsive web (Fig. 3).

As a method to verify the reliability of opinion mining of the platform, precision, recall and prediction accuracy experiments were performed. The data analyzed includes 2,452 positive Korean sentiment words and 2,000 positive English sentiment words on the basis of those sentiment words that are the unstructured text data for tourist information constructed through the development of this platform. Table 2 shows the results of analysis of the accuracy of opinion mining (OM) for the sentiment words in the platform, where the classes are divided into positive words and negative words.

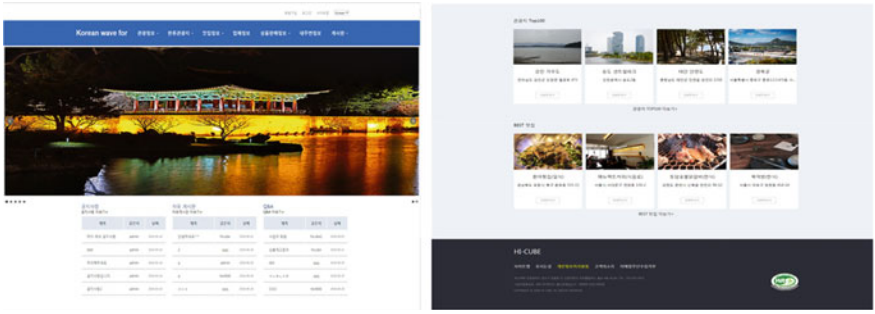


Fig. 3 Platform main screen

Table 2 Accuracy of opinion mining analysis

Detailed accuracy by class				
Class	Precision	Recall	F-measure	ROC area
Positive	1	0.932	0.964	0.982
Negative	1	0.912	0.953	0.982
AVG	*1*	*0.922*	0.958	*0.982*

The precision rates, which is the precision value, were shown to be 100% for both positive and negative words. The recall rates were 93.2% for positive words and 91.2% for negative words. The F-Measure values were 96.4% for positive words and 95.3% for negative words, and the average ROC curve showed the performance of 98.2 on average.

5 Conclusion

In this study, in order to revitalize the domestic tourism industry and respond to domestic and foreign tourists who need tourist information service, a visualization web platform that can provide more easily understandable and effective tourism information services was designed using the opinion mining analysis technique and the relevant platform was actually implemented. As a result, a conclusion that more objective and diverse contents can be provided to tourists who are users was obtained and though the foregoing, an infinite possibility of the spread of future-oriented one-to-one interactive tourist information service paradigms and service commercialization could be seen. In addition, to improve the image and recognition of the domestic tourism industry, which is currently in a stagnant period, using the attractive medium termed Korean Wave contents thereby leading successful shared growth of the domestic tourism industry and the Korean Wave contents industry, continuous

investments in tourism industry contents led by government-related organizations and research and development should be necessary.

References

1. Seo J, Jo H, Choi J (2015) Design for opinion dictionary of emotion applying rules for antonym of the korean grammar. JKIIT 13(2):109–117
2. Jo H, Seo J, Choi J (2015) OAR algorithm technology based on opinion mining utilizing stock news contents. JKIIT 13(3):111–119
3. Lee Y, Seo J, Choi J (2014) Fashion trend marketing prediction analysis based on opinion mining applying SNS text contents. JKIIT 12(12):168–170
4. Seo J-H, Lee H-S, Choi J-T (2015) Classification technique for filtering sentiment vocabularies for the enhancement of accuracy of opinion mining. Int J u- and e- Ser Sci Technol 8(10):11–20
5. Ghose A, Ipeirotis P, Sundararajan A (2007) Opinion mining using econometrics : a case study on reputation system. In: Proceedings of the 45th annual meeting of the association of computational linguistics, Prague, Czech Republic, pp 416–423

Analysis Design Study for Fake News Identification and Evaluation

Lee Won Park and Hangbae Chang

Abstract With the spread of the Internet and increasing amounts of self-proclaimed journalists, articles both true and inaccurate fill the web. These inaccurate articles, most commonly referred to as fake news, have proved to spread quickly and have immense social influence in society. Attempts to detect fake news articles through deep learning techniques and artificial intelligence prove the challenges in fake news detection. While detection techniques are still in development, there is not much research on how readers can discern fake news without technological aid. This paper addresses such limitations regarding the study of fake news detection and provide a detection model for readers. The model is based on logical steps built on detection cues mentioned in previous works. The appropriateness of the detection cues will be determined based on case studies.

Keywords Fake news · Fake news identification · Industrial security

1 Introduction

Fake news has always been around, and always has been a societal issue. However, the US presidential election in 2016 changed the public's perception of fake news. Previously, fake news was frequently referred to news parody or satirical news. The core distinction between fake news and parody or satirical news is that the former has the clear intention to deceive people, whereas the latter intends to make open jokes without the intention to provide false information. Additionally, previous works have proven that those who are familiar with this form of news parody or satire tend to be more politically informed and interested, compared to those who do not follow such commentary [1].

L. W. Park
Department of Security Convergence, Graduate School, Chung-Ang University, Seoul 06974, Korea

H. Chang (✉)
Department of Inudstrial Security, Chung-Ang University, Seoul 06974, Korea
e-mail: hbchang@cau.ac.kr

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_26

Fake news, however, can be defined as false news that aims to deceive readers and audiences. Fake news is meant to deceive, and thus often difficult to prove false. Although some believe that fake news is easily discernible, it has been proven that some people are especially susceptible to fake news because they strongly relate with topics of fake news. These groups of individuals also risk not being able to objectify the news they consume because they remain ingrained in their own social bubble, without finding the need to explore whether the news they consume is true or false [2]. However, certain social groups are not the only targets of fake news. Previous works have proven than individuals generally detect fake news correctly only half of the time—that is slightly above chance [3].

Previous works have made efforts to detect fake news based on the sentence structure, vocabulary, source of news, structure of news links, etc. Although the use of AI does aid the detection of fake news across the web, thus providing the potential to stop the initial spread of fake information, it does little to aid the detection process for individuals every day. Part of stopping the initial spread of fake news has to do with guiding individuals to also be able to detect fake news, and thus prevent them from sharing and thus participating in the spread of the fake news.

Therefore, this research is dedicated to exploring the methods by which people can easily detect fake news based on notable cues and logical steps. The research will present cues that will help people notice or at least doubt whether the news article is fake or not and observe whether the cues are applicable in real fake news articles.

2 Previous Works

Previous works surrounding fake news have often been centered around defining what fake news may incorporate, what forms fake news may take, and what fake news looks like in different forms of news media.

Previous works note that satire intentionally utilizes cues to reveal its own deceptiveness, since satirical news is not used as a method to deceive with malintent [4]. The research builds on previous work in satire detection to propose an SVM-based algorithm enhanced with 5 predictive features; grammar, humor, negative affect, absurdity, and punctuation. The paper also agrees with previous works in that people are generally unable to effectively recognize deception in news or in general circumstances.

Similarly, the problem arises when people consume fake news while unaware that it is fake news. Previous works argue that the most significant aspect about contemporary fake news is that it deliberately uses “empathic media”, news that targets the emotions of audiences. The paper notes that fake news can be created to build “fellow-feeling” or group emotion behavior within social networks”, and at heart fake news concerns “economics of emotion” [5].

Due to the fact that people will half the time or more, fail in discerning between fake and real news, previous studies have also made numerous attempts to provide insight on how to detect fake news via behavioral analysis, syntax, URL structures,

tone or wording of the article, etc. Previous works also provide 9 requirements for fake news detection, those being; the availability of both false and true instances, verifiability of "ground truth", digital textual format accessibility, homogeneity in lengths and in writing matter, manner of news delivery, predefined time frame, pragmatic concerns, and language and culture [6].

Other research suggests identifying reliable news sources, and assessing the pieces used by the journalist in the piece [7]. Additionally, the research suggests that people should be on the lookout for bias and logical fallacies. The paper classifies bias as something that can easily be identified through the words used in a piece. Since bias is subjective, words that express personal opinions or feelings may be indicative of bias in news articles. The research stresses the importance of looking for facts without any sort of non-objective descriptors.

As fake news is also often classified as clickbait, previous research has also made distinctions between content cues and non-text cues that may help discern clickbait [8]. The paper presents content cues as lexical and semantic levels of analysis, and syntactic and pragmatic levels of analysis. Non-text cues presented in the paper are image analysis and user behavior analysis. The paper notes that images can attract attention and because it can draw visual attention of readers, the image is processed before the article is read. Research notes that clickbait frequently uses images to draw in the interests of users through misinformation, and readers tend to receive spontaneously integrated information from the headline with a photo in the associated article.

3 Materials and Methods

3.1 Model

Based on previous works, this paper will present a logical method that will aid people in the process of discerning fake news online.

News consumers are advised to first consider the title of the news article. Previous works suggest that the title of the article may have different layers of fakeness to it. Such attributes of fake article titles may include (a) overrun sentences, (b) sentences with disparate topics, (c) sentences with unnecessary overuse of adjectives, (d) sentences that create strong emotions such as anger, disgust, hate, or sadness, and (e) topics that are completely novel.

After initially clicking on the article based on the title, readers will then be able to clearly note the source of the article. The URL and source may be quick and accurate benchmarks for readers to note whether the article is reliable or not. Once readers have checked the title and source of the article, they will then dissect the body of the article. The body of the article is analyzed similarly to how the title is analyzed. Since there is more text in the body than the title, there may be more frequent and obvious non-textual and textual cues that indicate the level of fakeness

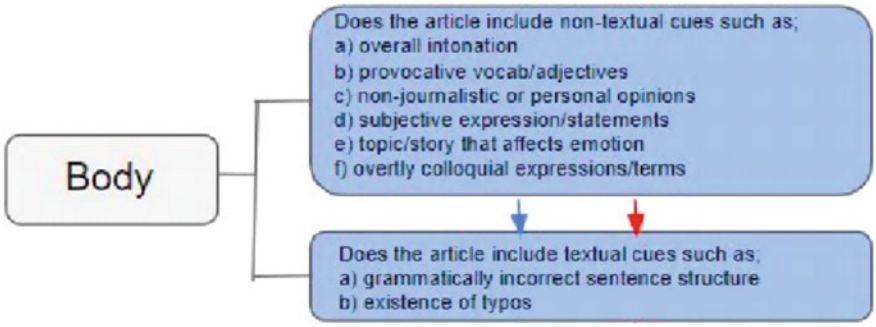


Fig. 1 Discerning fakeness from the body of the article

of the news article. Non-textual cues may include the intonation of the article, the level of humor or satire, the use of inflammatory adjectives, the use of subjective descriptions, the existence of opinion of the writer, etc. Textual cues may include the sentence structure of the article, the grammatical accuracy of the sentences within the article, the existence or frequency of typos, etc.

Finally, readers may be able to discern the level of fakeness of a news article based on the inserted photo in the article. Manipulated photos are of course fake news, but the article can also be suspected as a fake if the inserted photo is completely irrelevant to the title or context of the article (Fig. 1).

Based on the assessment points discussed above, a logical algorithm was developed. However, this paper will only examine and test fake news identification cues via the body of text. The logical algorithm for analyzing the body of text is presented below.

3.2 Case Study Test

Based on the logical detection model suggested above, this paper tests how appropriate and applicable the model is on existing fake news articles. The article presented is primarily judged on the body of the article. The fake news article introduced in the paper is one of the top fake news articles that received the most hits of Facebook in 2018. The article title is "Muslim Figure: "We have Pork-Free Menus Or We Will Leave U.S." How Would You Respond This?" The text of the article is presented below in italics.

"Muslim migrants have turned into a worldwide issue, and the circumstance deteriorates each day. Our nation has had a considerable measure of Muslim-related issues of late, yet this time settlers went too far. President Donald Trump had the ideal answer for the issue forced by outsiders, however Democrats didn't bolster him. You won't think about what occurred next…

Table 1 Identified cues for fake news article

Cues	
Grammatical error	✓
Typo	
Overall undertone	✓
Provocative adjectives/statements	
Non-journalistic/personal statements	
Subjective expressions/descriptions	✓
Novel/emotional topic	
Colloquial terms/statements	

Barack Obama urged Muslims to request their rights. Muslims were informed that America is an astonishing nation loaded with potential outcomes. All things considered, it wasn't some time before settlers set particular prerequisites. Some of these individuals do have issues, and need genuine help. Be that as it may, the vast majority of them are simply crossing the US outskirt to appreciate greatest advantages.

President Trump realized this would transform into a difficult issue, and recommended that our nation bars foreigners from a few Muslim dominant part nations. Tragically, the travel boycott was blocked comfortable start, and Muslim outsiders attacked the nation (Table 1).

America resembles a strong and minding host, however some of its guests are accompanying lethal thoughts in their psyche. They are continually requesting more. This time Muslim settlers made a particular request, and it includes sustenance served in schools."

The analysis of the fake news article is as follows:

- "Issues of late" is grammatically incorrect [textual cues: grammatical error]
- The term "ideal answer" is subjective [non-textual cues: subjective expression]
- The term "difficult issue" is subjective [non-textual cues: subjective expression]
- The phrase "the travel boycott was blocked comfortable start" is grammatically questionable [textual cues: grammatical error]
- The term "attacked" is both provocative and non-factual [non-textual cues: provocative expression and subjective expression]
- The term "America resembles a strong and minding host" uses subjective descriptive terms [non-textual cues: subjective expression]
- The general undertone of the article is politically provocative, resembles hate speech due to subjective and emotional choice of vocabulary, and attempts to make political statements by introducing particular political figures [non-textual cues: provocative overall undertone of article]

4 Conclusion

The research presented in this paper contributes to the current state of fake news research by providing insight on how readers can initially spot fake news to discern or suspect the integrity of a news article. Technological advancements are key to stopping the rapid spread of fake news, individuals are unable to effectively utilize such technologies every day. This paper recognizes the limitations of previous studies and presents a logical map that will help readers to discern or suspect whether the article they are consuming is fake or true.

Only one fake news article was introduced and analyzed in this paper. The article was selected to demonstrate the typical form of fake news that is difficult to discern at first glance due to its length, subject matter, and seemingly legitimate journalistic format. Future research should be conducted on an extended level with more sample articles provided for analysis.

References

1. Tandoc E Jr, Lim Z, Ling R (2018) Defining "Fake News". Digi Journalism 6(2):137–153
2. Spohr D (2017) Fake news and ideological polarization: filter bubbles and selective exposure on social media. Bus Inform Rev 34(3):150–160
3. Frank MG, Feeley TH, Paolantonio N, Servoss TJ (2004) Individual and small group accuracy in judging truthful & deceptive communication. Group Decis Negot 13:45–59
4. Rubin, V., Conroy, N., Chen, Y., & Cornwell, S.: Fake News or Truth? Using Satirical Cues to Detect Potentially Misleading News. Proceedings of NAACL-HLT, 7–17. (2016)
5. Bakir V, McStay A (2018) Fake news and the economy of emotions. Digit Journalism 6(2):154–175
6. Rubin V, Conroy N, Chen Y (2015) Towards news verification: deception detection methods for news discourse. https://doi.org/10.13140/2.1.4822.8166
7. Ireland S (2018) Fake news alerts: Teaching news literacy skills in a meme world. Refer Librarian 59(3):122–128
8. Chen Y, Conroy N, Rubin V (2015) Misleading online content: recognizing Clickbait as "False News". ACM Workshop on Deception Detection

System Architecture with Respect to Indoor Localization

Martin Zmítko, Jan Budina, and Tomáš Kozel

Abstract The aim of this paper is to analyze basic system architectures and databases for mobile applications specialized in indoor localization. First, it describes system architecture for indoor localization and databases as the bases of all system. Next, several types of system architectures are explored and it is discussed, which of them is the most suitable for indoor localization usage. The paper also describes the most widely used commercial system for localization. The goals of this paper to define and compare different types of system architecture, then compare special needs for indoor localization and review commercially used systems.

Keywords Fingerprint · Beacons · Indoor localization · Architecture · System · Database · Analysis · Mobile application

1 Introduction

Nowadays, the mobile platform of computing devices is starting to dominate over the static one. Many statistics [1] show that there was the overturn came in 2014, and since then mobile devices have been beating desktop ones as for site traffic. Mobile devices often work with web pages designed for mobile devices, but this constitutes certain limitations to the basic principles of this technology such as client-server communications [2]. Software developers respond to this problem with mobile apps that are installed on mobile devices and provide much greater added value to the user. This is mainly due to the possibility of automatic synchronization, options for using dynamic content and, last but not least, the ability to quantify all information from the device and its sensors that can improve application performance. Basic iBeacon device for indoor localization purposes without application is useless. If we want to use iBeacon for indoor localization, it is necessary to develop a complex system [3]. This paper rings a new view of present approaches in relation to mobile and system

M. Zmítko (✉) · J. Budina · T. Kozel
Faculty of Informatics and Management, Department of Informatics and Quantitative Methods, The University of Hradec Kralove, Hradec Kralove, Czechia
e-mail: martin.zmitko@uhk.cz

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_27

architectures. The purpose of this is to inform the reader about the areas of architecture approach to indoor localization [4, 5]. Looking closer at both approaches, we see that the two worlds are gradually beginning to mingle. Embedded Web technologies enhance the capabilities of integrated applications, which can conveniently transmit dynamic content, read data from internal phone sensors, and are platform-independent. Of course, this technology does not meet all the requirements that application developers have. On the other hand, the mobile application is platform-dependent. It is therefore necessary to have applications separate for the two largest mobile worlds. The application, however, brings many benefits in the form of almost unlimited individualization that leaves open multiple possibilities for developers [6, 7].

2 Technology Review

Due to the use of indoor localization, it is necessary to use mobile devices that form the basis of the entire system. It brings some specifics in the form of using sensors and processing measured data. The mobile platform has some special features and requirements to focus on [8], C**onnectivity with the environment**—devices interact with nearby devices and enable data exchange; **Using sensors**—collecting information from built-in phone sensors for application usage; **Approaches to mobile System architecture**—method of representation the mobile system devices; it can be a native application or a responsive website or a hybrid scenario; **Secure communication**—necessary security of the communication channel, due to the fact that a large amount of information is sent between the mobile device and server part; **Usable user interface**—user friendly interface, support gestures, forms, **Energy saving**—optimizing the code for the lowest power consumption in the active state.

These characteristics are typical for the mobile system architecture and must be taken into account during application development in (Table 1).

3 Related Works

The following chapter describes an existing solution with examples of different types of system architecture approaches. Each of the applications is briefly described with a specific architecture and used technologies.

Table 1 Native application and responsive webpage comparison

	Native application	Responsive webpage
Portability	Based on system	Cross-platform
Sensor availability	All	Limited: Geoloc, accel,
Performance	+	–
User-friendly	+	–
Functionality	+	–
Security	+	–
Programming difficulty	–	+
Serviceable	–	+
Navigation	+	–
Upgradability	–	+
Findability	–	+
Shareability	–	+

3.1 Smart Campus

Architecture: Native mobile applications, .NET backend with NoSQL databases.

The first example is Smart Campus. It is an application consisting of a mobile and server part focused on indoor localization using WiFi signals. Backend consists of two parts: **RadioMapService3**—Part containing a radio map and metadata about the building. **WifiSnifferPositioningService**—A position finding service that needs to be deployed on your own application server. Information is stored in the NoSQL database (Couchbase), which seems to be an appropriate solution for large amounts of data [9, 10]. The mobile part is represented as a native application for the Android and iOS platforms communicating with the backend part using the RESTful interface and JSON format. Data storage is used to store local data in the form of local data or SQLite until the server part is synchronized [11].

3.2 Rainforest of Americas

Architecture: Native mobile application, NoSQL database.

The second example is the native Rainforest of the Americas [12]. This application serves the visitors of the zoological and botanical gardens of Los Angeles as support in their field excursion. This is an application that provides extra-standard exposure information based on the context. The data obtained serves as feedback to the zoo owner regarding the movement of people and the attendance of individual sections of the garden. From the technical point of view, this is a native application using

Cubecon cloud service backend [13]. This service also uses third-party solutions in the form of a database provided by Mesosfer [14] that provides data stored on their cloud based on NoSQL databases. The Mesosfer service itself supports a number of platforms. From mobile to native application to the web interface and responsive design. In this case, we see the use of several technologies. With such extensive cloud storage services for IoT devices (including iBeacons), using NoSQL is a necessary step.

4 Comparison

We choose the most widespread solution based on our research and existing works. Each approach is described in its own section with its pros and cons.

A—Native Mobile Applications as a Thick Client

The entire application is built as one large package, containing all data and information the user might need.

B—Native Mobile Application with Server Backend

In this case, we are building a native application (Android, iOS, ..) that should communicate with the server component (using the REST API-application program interface that uses HTTP requests to GET, PUT, POST, DELETE data). Against the first variant, there is a rapid thinning of the application portion. The main advantage of this approach is the already mentioned smaller size of the application itself. However, if we move the data to the server site, data transmission grows and a better connection quality is required. This architecture provides a possibility of extending the server part in a variety of aspects. Among others, it can be the integration of other systems, connection to a variety of APIs, or storage of custom data into databases on application servers.

An example of this approach can be a zoo that stores the most basic data in a native application (for example, a sitemap, schedule, opening hours). The remainder of the dynamic data is then provided by the server portion upon request from the application client according to the user's location.

C—Hybrid Application with the Server Backend

This approach is used, for example, by Cordova plugin. It is essentially a responsive web application with a server part. This can be seen as a web page interacting with the user and containing relevant data. In this case, the mobile application is only a wrapper—the next layer (which is created by application plugin) that provides an interface between the Web part and the native features of the device (no matter whether it comes to the Android or iOS functionality). The architecture has to deal with worse support for browsers (or HTML5, JavaScript libraries) and their access to mobile device sensors. This approach allows you to install an application on a specific

device, which, however, requires higher implementation requirements (multiplatform approach). This approach can be used when you want to give the user the widest possible use of the upcoming application (from the perspective of operational system). The web part is available without the need to install an application, but positioning functionality is currently limited.

D—Responsive Web Application with Server Backend

System for indoor localization purely on web technologies in cooperation with the server part.

5 Conclusions

The aim of this section is to find the most appropriate approach to creating indoor localization for a specific use case. Specifically, it is applied to the University of Hradec Králové—a school application for the support of education, which would be used by students and teachers as their aid. It is expected that the application will be used by up to 2,400 users. The emphasis is placed on the possibility of linking the upcoming solution to other systems, the speed of providing materials for education, the complexity of implementation and the overall friendliness of the whole solution.

Each architecture was designed during the experiment and the following Table 2 shows the comparison of individual architectures from the point of view of the individual criteria that are relevant to designing the system. Each criterion belongs to a metric that determines its importance in architecture.

metric: 1–10 (10 highest priority, 1 least important)

rating: 1–5 (5 best, 1 worst)

Table 2 Defined approach comparison

	Metric	A	B	C	D
Application size	4	2	3	4	5
Data consumption	6	1	2	3	2
System compatibility	10	3	3	3	5
Offline work	2	5	4	3	1
Localization accuracy	9	4	4	4	3
Update speed	3	2	2	3	4
Backend severity	8	5	3	3	3
User experience	7	4	5	4	4
Security	5	4	4	3	3
Implementation difficulty	1	2	3	3	3
Evaluation		18,6	18,6	18,5	**19,3**

Based on a weighted average rating, best-performing D has the best rating.

It is clear from the above paragraphs that not all architectures can be universally used. Now we know that it is possible to build a thick native application, mobile application with a server part, hybrid application or use web technologies. It always depends on the specific case, where we want to operate the indoor localization. There are cases when indoor localization is used only in buildings, and other cases when it is used both in buildings and open spaces outside the buildings. If we consider this hybrid scenario that combines inside and outside, we cannot properly decide, which solution is the best and we have to look for a compromise between all suggested solutions. If we want to navigate users in a closed building covered by a Wi-Fi signal, the last "D" option is ideal because we can quickly distribute updates to all devices that do not mind replacing large amounts of data. Otherwise, when locating occurs in locations with poor or no data coverage, it is ideal to choose the first option, where the entire application with all the data is stored on the mobile device. This gives users all the data without having to connect to the data network. However, there are limitations in the form of less frequent updates and the need to create separate applications for different mobile platforms.

"C" option would be a compromise, which would use both technologies, a mobile application while integrating a web browser. The advantage is that when you connect to a mobile network, data can be downloaded directly from the server, and offline data on the mobile device would be used where no data is available.

Conflicts of Interest and Funding Statement

The authors declare that there is no conflict of interest regarding the publication of this paper. This work was supported by the SPEV project, financed from the Faculty of Informatics and Management, University of Hradec Kralove.

References

1. Chaffey D (2018) Mobile marketing statistics. Mobile Market Statist Comp, led. 21, 2018. https://www.smartinsights.com/mobile-marketing/mobile-marketing-analytics/mobile-marketing-statistics
2. Wang AI, Jarrett M, Sorteberg AE (2009) Experiences from implementing a mobile multiplayer real-time game for wireless networks with high latency. Int J Comput Games Technol, roč, 1–14. https://doi.org/10.1155/2009/530367
3. Mora H, Mora Gimeno FJ, Signes-Pont MT, Volckaert AB (2019) Multilayer architecture model for mobile cloud computing paradigm. Complexity, roč, 1–13. https://doi.org/10.1155/2019/3951495
4. Budina J, Klapka O, Kozel T, Zmitko AM (2015) Method of iBeacon optimal distribution for indoor localization. In: Modeling and using context, roč. 9405, H. Christiansen, I. Stojanovic, a G. A. Papadopoulos, Ed. Cham: Springer International Publishing, pp 105–117
5. Kriz P, Maly F, Kozel AT (2016) Improving indoor localization using bluetooth low energy beacons. Mob Inf Syst, roč, 1–11. https://doi.org/10.1155/2016/2083094
6. Budina J, Zmitko AM (2018) Idea of learning context system using iBeacons. 2018 ELEKTRO, Mikulov, kvě, 1–6. https://doi.org/10.1109/elektro.2018.8398378

7. Sethi a P, Sarangi SR (2017) Internet of things: architectures, protocols, and applications. J Electr Comput Eng, roč, 1–25. https://doi.org/10.1155/2017/9324035
8. Wasserman AI (2010) Software engineering issues for mobile application development. In: Proceedings of the FSE/SDP workshop on Future of software engineering research—FoSER '10, Santa Fe, New Mexico, p 397. https://doi.org/10.1145/1882362.1882443
9. Mongo DB (2018) MongoDB architecture, led. 11. https://www.mongodb.com/mongodb-architecture
10. SmartCampusAAU. https://github.com/BentThomsen/SmartCampusAAU
11. Hansesn R, Thomsen AB (2018) SmartCampusAAU—an open platform enabling indoor positioning and navigation. SmartCampusAAU—an open platform enabling indoor positioning and navigation, led. 30, 2018. https://www.semanticscholar.org/paper/SmartCampusAAU-An-Open-Platform-Enabling-Indoor-Po-Hansen-Thomsen/be26b716fdd80e9f35af18ccbff15eac6c873847
12. Girish D (2018) How Beacons can help Tansform Zoo and Theme Park Experience for Visitors, How Beacons can help Tansform Zoo and Theme Park Experience for Visitors, led. 30, 2018. https://blog.beaconstac.com/2014/09/how-beacons-can-help-tansform-zoo-and-theme-park-experience-for-visitors
13. Avianto (2018) How Your Business can benefit from Usage of Cubeacon", How Your Business can benefit from Usage of Cubeacon, led. 30. https://blog.cubeacon.com/how-your-business-can-benefit-from-usage-of-cubeacon.html
14. Mesosfer, led. 30, 2018. https://mesosfer.com/feature.html

Deep Learning Adoption Blockchain Secure Framework for Cyber Physical System

Jose Costa Sapalo Sicato, Jeonghun Cha, Yi Pan, and Jong Hyuk Park

Abstract The growth of cyber-physical systems (CPS) has greatly improved industrial efficiency and consumer convenience. Nevertheless, the widespread deployment of 5G-CPS systems also poses new problems that are of major concern to security. In order to help create stable and robust 5G-CPS systems, systematic analyzes are carried out on CPS and IoT from individual smart devices that act as terminal nodes to the entire network. The results of the review will help to find weaknesses in CPS systems and/or provide guidance on how to create protection in these systems using deep learning and blockchain techniques. Risk reduction approaches that can help improve protection will also be addressed.

Keywords CPS · Deep learning · Blockchain · Security and privacy

1 Introduction

Cyber-physical systems (CPS) are defined as engineered systems based on the close integration of cyber entities (e.g. computation, communication, and control) and physical objects (natural and man-made structures controlled by the laws of physics). CPS is known to be the adult production of IoT, the completion of the IoT definition and vision. These are combined into a closed-loop, offering design frameworks and understanding all aspects of networked composite systems that are monitored and

J. C. S. Sicato · J. Cha · J. H. Park (✉)
Department of Computer Science and Engineering, Seoul National Univerisy of Science and Technology, Seoul 01811, Korea
e-mail: jhpark@seoultech.ac.kr

J. C. S. Sicato
e-mail: josecostasicata@seoultech.ac.kr

J. Cha
e-mail: ckwjdgns@seoultech.ac.kr

Y. Pan
Department of Computer Science, Georgia State University, Atlanta, Georgia
e-mail: yipan@gsu.edu

J. J. Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 716, https://doi.org/10.1007/978-981-15-9309-3_28

managed by computer algorithms and are closely interlinked between users and the Internet. That is, hardware and software organizations are interconnected and usually work on various time and location-based scales. In reality, IoT (through sensors and actuators) enables the interaction between the virtual world and the physical world [1]. CPS consists of a great number of devices, which is considered as a recent paradigm that its concept involves a large number of interconnected devices based on the presence of objects such as sensors, actuators, mobile devices, Radio Frequency Identification (RFID) tags. The internet of things purpose is to make intelligent devices around us to transform our way of living today and perform the daily tasks. Everyday using things like smart transportation, smart watches, smart phones etc., Connect to the internet so critical information can be sent over the network anywhere and anytime [2].

With the increasing number of CPS devices traditional 4G wireless technology cannot meet the speed, latency, efficiency requirements of CPS. Therefore involving the fifth generation (5G) network, which is becoming more available of the growth of the IoT applications [3]. Due to the massive connections in huge number of CPS devices, its required performance criteria such as security, ultra-low latency, coverage of wireless communication, connectivity [4]. To achieve these challenges, enabled 5G will connect through device-to-device D2D, and various technologies are applied such as massive multiple-input output (MIMO), Blockchain, Deep Learning will make contribution to meet the demand for massive wireless service providers to simulate new social and economic development [5]. The development of 5G CPS technology is key to better quality of life than ever before, however, the security risks are becoming ever more acute [5]. What's more, CPS-IoT Has difficulty assessing interaction-related threats and vulnerabilities, and new security issues emerge.

Therefore in this paper we presented solutions for 5G-CPS and it's organized as follows: Sect. 2 describes the related work and CPS challenges. In Sect. 3 we proposed solutions using Deep learning adoption blockchain for 5G CPS and finally Sect. 4 conclude this paper.

2 Related Work

In cyber physical system, security and privacy are the main challenges, prospective measures based many research efforts highlight the CPS security issues, and mentioned that CPS technology operates in three main layer application layer, network layer and perception layer. Many attacks can be performed in IoT. It was mentioned that vulnerabilities of IoT devices to malware attacks and risk of physical compromise of devices poses significant danger to IoT support. Due to the emerging attacks embedded devices, it face limitations in terms of computational power, connectivity and energy budget.

2.1 CPS Challenges

According to CPS framework we focus on concerns on some specific issues based on CPS, which we present any emerging technology.

- Heterogeneity: The heterogeneity stems a wide variety of CPS implementations and the different issues concerned. Such aspects of heterogeneous in terms of computational complexity, communication processes, versatility, storage space, energy consumption, adaptability, longevity and variable data size [6].
- Security: Security and privacy come into play in a range of applications. In nearly every sector, which includes a cyber-component, there are reports of cyber-attacks daily. In order to ensure a comprehensive architecture of IoT environment, data privacy and confidentiality of information security must be properly handled [7].
- Trust: which is essential to the implementation of the cyber physical system. Nonetheless, the primary concern about loT protection and privacy is the flow of data to the network that will be exchanged and used by various applications (service providers).
- Scalability: In an important concept in design of a future poor CPS network architecture to support the potential growth in the number of applications and the amount of information that they generate. Requirements for CPS are anticipated in a variety of areas: manufacturing, smart agriculture, smart logistics, smart transport, smart grid, smart environmental protection, smart defense, healthcare, smart home, and others, which it gives a implication of the high scalability issues. New architectural approaches are needed while keeping the problem of manageability in mind [8].
- Latency: Delay occurs after the input port of the device reads the data, but before the output port sends the new data, while the transmission time and the transmission latency occurs after the output port of the front end. The part sends the data but before the next input port Component is reading the data.

2.2 Key Considerations

Considering the characteristics of a 5G CPS network's security, it concerns confidentiality, integrity, availability, reliability, traffic security, authentication and protection. Therefore 5G-CPS needed security features require the following key requirements mechanism:

- Authentication—To ensure devices connected to 5G CPS are reliable as they intended to be, is required a strong authentication. Device should have authenticated to the network for any IoT implementation in CPS-IoT.
- Privacy—To make sure that all data is only transmitted between approved systems and devices. A security camera, for example, should send information to its owner's smartphone, and vice versa.

- Confidentiality—some of the data sent to "things" are proprietary and are to be protected. A proper encryption procedure is required to ensure data confidentiality as IoT data travels through several network hops [9]. Owing to the dynamic integration of software, devices, and networks, data stored in IoT devices are vulnerable to privacy infringements by breaching existing IoT network nodes.
- Integrity—IoT integrity should take into consideration the avoidance, identification or blockage of network attacks on information shared between sensors and actuators or controllers. Therefore verify that message remains unchanged or has been modified.
- Reliability—CPS has been deployed and is intended to operate for a long time, and as several CPSs provides availability requirements, it is very challenging to upgrade them or correct their faults. It is therefore important to include the Maintenance of these processes reactive and predictive. Reactive maintenance makes it easy to identify and repair defective devices with the aid of an IoT monitoring system that sends notification of any discovered problems [10].

3 Proposed Framework

Here, we are proposing a secure and intelligent framework for the effective sharing of data with different IoT facilities for the different layers. The whole deep learning and blockchain-based architecture is shown in Fig. 1.

The proposed framework consists of 3 layers where traffic flows through sensors and physical layer devices are first sent to network layer using 5G massive MIMO and routers devices; this device generates a hash of edge node using standard hash

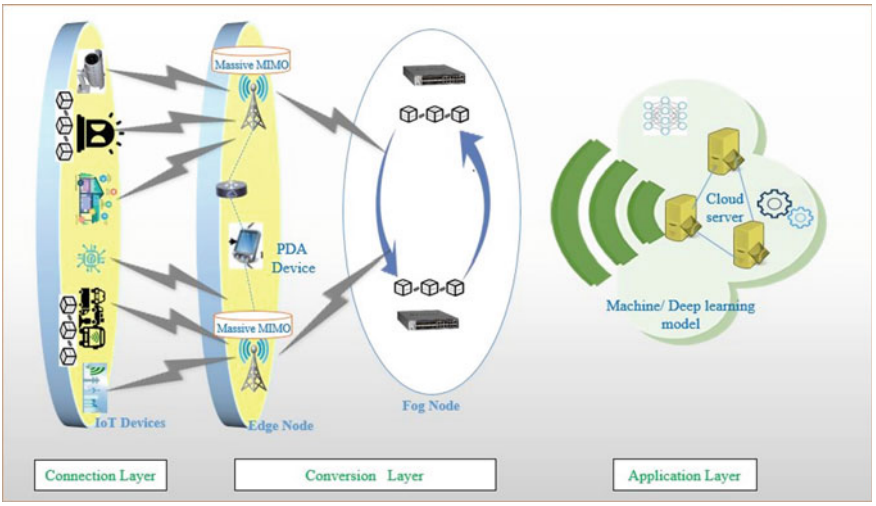


Fig. 1 DL Adoption Blockchain for CPS

algorithms and then forwards the hash to a private blockchain network. Blockchain is a distributed and tamper-resistant ledger that does not rely on a centralized authority to create trust with a core layer framework for decentralized trust management. Blockchain technology will ensure that the network is tamper proof. In fog node the latency decreases by processing the fog node data obtained from incoming CPS traffic. The intercommunication connectivity between connected devices and security is provided with blockchain technology.

In the blockchain, any facility, such as smart city, smart home, and others, will serve as a blockchain node. The hash sent from the edge node system will be received by each node and the data block must be validated and verified by the nodes. Verification is performed on the basis of the received hash and this hash is compared to the hash of the previously received data block. Some of the blockchain nodes need to validate the block.

The PDA system in fog node encrypts the real traffic data that use the key and encrypted data, the fog node data hash, the ID is submitted to the cloud-based database server. Whether someone is trying to tamper with it block of data, the next blocks can be also be affected. This is given as feedback to various machine learning-based various applications. We addressed earlier the different deep applications training strategies for various CPS uses, such as predictive smart grid, smart healthcare. The key benefit of our proposed system is the safe exchange of data between different node layer facilities and the data can be used for different IoT applications.

4 Conclusion

Cyber physical system is attracting many researchers to carry out research in this field. In our paper was provided any overview of different challenges that occur in CPS devices and which can compromise user security and privacy. Based on these issues and challenges was proposed solutions using deep learning adoption blockchain secure framework for CPS applications.

Acknowledgements This study was supported by the Advanced Research Project funded by the SeoulTech (Seoul National University of Science and Technology).

References

1. Shih CS, Chou JJ, Reijers N, Kuo TW (2016) Designing CPS/IoT applications for smart buildings and cities. IET Cyber-Phys Syst: Theory Appl 1(1):3–12
2. Sonar K, Upadhyay H (2014) A survey: DDOS attack on internet of things. Int J Eng Res Dev 10(11):58–63
3. I-Scoop, 5 g and iot in 2018 and beyond: the mobile broadband future of iot,[Available on line 14 Jan 2018], https://www.i-scoop.eu/internet-of-things-guide/5g-iot/

4. Jaiswal N, Analysys mason, 5g: continuous evolution leads to quantum shift, [Available on line 14 Jan 2018]. https://www.telecomasia.net/content/5gcontinuous-evolution-leads-quantum-shift
5. Hatzivasilis G, Fysarakis K, Soultatos O, Askoxylakis I, Papaefstathiou I, Demetriou G (2018) The industrial internet of things as an enabler for a circular economy Hy-LP: a Novel IIoT protocol, evaluated on a wind park's SDN/NFV-enabled 5G industrial network. Comput Commun 1(119):127–137
6. Jirkovský V, Obitko M, Mařík V (2016) Understanding data heterogeneity in the context of cyber-physical systems integration. IEEE Trans Industr Inform 13(2):660–667
7. Giraldo J, Sarkar E, Cardenas AA, Maniatakos M, Kantarcioglu M (2017) Security and privacy in cyber-physical systems: A survey of surveys. IEEE Design Test. 34(4):7–17
8. Huegel Richa C, Lucena MM, Passig Horstmann L, Conradi Hoffmann JL, Fröhlich AA (2020) Modeling Time requirements of CPS in wireless networks. Sensors 20(7):1818
9. Aref MA, Jayaweera SK, Machuzak S (2017) Multi-agent reinforcement learning based cognitive anti jamming. In: Proceedings IEEE wireless communication and networking conference (WCNC), pp 1–6, San Francisco, CA
10. Peng H et al (2019) Reliability analysis of CPS systems under different edge repairing strategies. Physica A: Statist Mech Appl 532:121865