Lecture Notes in Electrical Engineering 590

James J. Park
Laurence T. Yang
Young-Sik Jeong
Fei Hao *Editors*

Advanced Multimedia and Ubiquitous Engineering

MUE/FutureTech 2019

Springer

Lecture Notes in Electrical Engineering

Volume 590

Series Editors

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

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

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

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

China

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

India

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

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

Japan

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

South Korea

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

Southeast Asia

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

USA, Canada:

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

All other Countries:

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

Christoph Baumann, Executive Editor (christoph.baumann@springer.com)

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

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

James J. Park · Laurence T. Yang ·
Young-Sik Jeong · Fei Hao
Editors

Advanced Multimedia and Ubiquitous Engineering

MUE/FutureTech 2019

Springer

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

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

Laurence T. Yang
Department of Computer Science
St. Francis Xavier University
Antigonish, NS, Canada

Fei Hao
School of Computer Science
Shaanxi Normal University
Xi'an, China

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-32-9243-7 ISBN 978-981-32-9244-4 (eBook)
https://doi.org/10.1007/978-981-32-9244-4

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

Contents

MUE2019

Near Infrared, Long-Wave Infrared and Visible Image Fusion Based on Oversampled Graph Filter Banks

YuLong Qiao[1], XueYing Gao[1(✉)], and ChunYan Song[2]

[1] College of Information and Communications Engineering, Harbin Engineering University, 145 Nantong Street, Nangang District, Harbin 150001, Heilongjiang, People's Republic of China
943377654@qq.com

[2] College of Mechanical and Electrical Engineering, Northeast Forestry University, 26, Hexing Road, Dongli District, Harbin 150040, Heilongjiang, People's Republic of China

Abstract. The near infrared (NIR), long-wave infrared (LWIR) and visible image fusion combines the interesting information of NIR and LWIR images and the details of visible images, which can provide an informative image for human visual perception and subsequent processing. In this paper, we propose a novel image fusion method based on oversampled graph filter banks. And then, the fusion rules for low frequency subbands and high frequency subbands are constructed by a new method of autoselect weighted average or maximum absolute value independently. Finally, the fusion image is obtained by applying the inverse transform on the fusion subband. The experimental results show that the proposed method outperforms the recently proposed 4 methods in terms of visual effects and performance indicators.

Keywords: Graph signal processing · Image fusion · Graph filter banks · Oversampled

1 Introduction

Infrared and visible image fusion is an important part of multi-sensor image fusion technology. Infrared light can be divided into near infrared (NIR), mid-wave infrared and long-wave infrared (LWIR) according to wavelength. At present, there are many NIR and visible image fusion [1] or LWIR and visible image fusion [2–4], but there is little research on NIR, LWIR and visible image fusion. NIR electromagnetic waves with wavelengths ranging from 0.78 to 1.5 μ are very sensitive to plants, and almost all plants can completely reflect them. LWIR wavelength is 1.5 to 400 μ, which is mostly absorbed by surface skin. According to the imaging principle of infrared image, even in the case of low visibility, the target can still be easily found through infrared image. Compared with the infrared image, the visible image has stronger contrast and can reflect more details. In order to effectively synthesize the scene information of NIR, LWIR and visible images, a multi-scale decomposition method based on oversampled graph filter banks [5] is proposed in this paper. Because it is not only similar to the

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 3–10, 2020.
https://doi.org/10.1007/978-981-32-9244-4_1

traditional multi-scale method, but also has good ability to decompose rules or irregular signals, it can effectively preserve the edge and texture information of the source image, so we use the oversampled graph filter banks to decompose the source image. Although current fusion methods based on simple fusion rules (such as average and weighted average, pixel absolute value, and regional characteristics, such as regional energy, entropy, variance, average gradient and contrast) can be extended to multi-image fusion, but its fusion effect is very limited. At the same time, the fusion rule models with better fusion effect are only suitable for the fusion of two images, and can not be extended to multi-image fusion. For this reason, a new fusion method suitable for multi-image is constructed in this paper.

The overall structure of this paper is summarized as follows: The second part mainly introduces the simple principle knowledge of oversampled graph filter banks. The new NIR, LWIR and visible image fusion method we proposed is described in the third part. The fourth part introduces the experimental parameter setting, experiment comparison and result analysis. We get the conclusion in the last part.

2 Basic Theory

In this paper, we only consider a finite undirected graph with no-loops or multiple links. Generally, a graph G can be represented as $G = \{V, \varepsilon\}$, where V and ε denote sets of nodes and edges, respectively [6]. The number of nodes is $N = |V|$, unless otherwise specified. The $(m, n) - th$ element of the adjacency matrix $\mathbf{A}$ is defined as follows:

$$\mathbf{A}(m,n) = \begin{cases} \omega_{mn} & \text{if nodes } m \text{ and } n \text{ are connected} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where ω_{mn} is the weight of the edge between m and n. The graph Laplacian matrix and the symmetric normalized version are defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$ and $\mathbf{L}' = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$, where the degree matrix $\mathbf{D}$ is a diagonal matrix. Sakiyama et al. [5] described a method of oversampling signals defined on a weighted graph by using an oversampled graph Laplacian matrix. Let us define the original graph Laplacian matrix of a bipartite graph as $\mathbf{L_0}$, and its corresponding adjacency matrix $\mathbf{A_0}$ whose size is $N_0 \times N_0$. Without loss of generality, an oversampled symmetric normalized graph Laplacian matrix can be represented as $\tilde{\mathbf{L}}' = \tilde{\mathbf{D}}^{-1/2}\tilde{\mathbf{L}}\tilde{\mathbf{D}}^{-1/2}$, where

$$\tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{A}_0 & \mathbf{A}_{01} \\ \mathbf{A}_{01}^T & \mathbf{0}_{N_1 - N_0} \end{bmatrix} \tag{2}$$

Where $\mathbf{A}_{01}$ contains information on the connection between the original graph and the appended nodes so that $\tilde{\mathbf{L}}$ is still a bipartite graph. And than we define the original signal as f_0, the signal for additional nodes as f_1 whose length is $N_1 - N_0$. So the oversampled signal is represented as $\tilde{f} = [f_0 \quad f_1]^T$.

Since the expansion process of the signal and its corresponding oversampled symmetric normalized graph Laplacian matrix is independent of filter selection, the entire transformation is perfect reconstruction only when the graphics filter bank satisfies the perfect reconstruction conditions. In this paper, we use two-channel oversampled graph filter banks as shown in Fig. 1. So the whole transform of oversampled graph filter banks as follows:

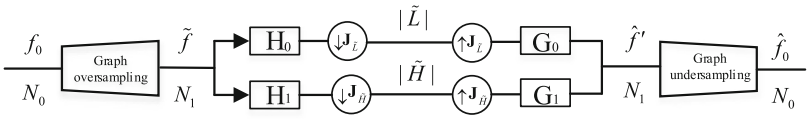


Fig. 1. Graph oversampling followed by two-channel oversampled graph filter banks.

$$\begin{aligned}\mathbf{T} &= \frac{1}{2}\mathbf{G}_0(\mathbf{I}-\mathbf{J})\mathbf{H}_0 + \frac{1}{2}\mathbf{G}_1(\mathbf{I}+\mathbf{J})\mathbf{H}_1 \\ &= \frac{1}{2}(\mathbf{G}_0\mathbf{H}_0 + \mathbf{G}_1\mathbf{H}_1) + \frac{1}{2}(\mathbf{G}_1\mathbf{J}\mathbf{H}_1 - \mathbf{G}_0\mathbf{J}\mathbf{H}_0) = \mathbf{I}_N\end{aligned} \tag{3}$$

The filtering in the spectral domain is defined also similarly to the critically sampled case [6]: $\mathbf{H}_k = \sum_{\lambda_i \in (\tilde{G})} h_k(\lambda_i)\tilde{\mathbf{P}}_{\lambda_i}$ and $\mathbf{G}_k = \sum_{\lambda_i \in (\tilde{G})} g_k(\lambda_i)\tilde{\mathbf{P}}_{\lambda_i}$,

where $\tilde{\mathbf{P}}_{\lambda_i}$ is the projection matrix [5] of the oversampling graph.

3 Proposed Method

An image is a regular 2-D signal and it can be viewed as a graph signal by connecting each pixel with its neighboring pixels (edges) and interpreting a pixel values as the function value on a node. In the work [7], they use 8-connected graph representation of an image shown in Fig. 2(a). In order to do "separable" transform for an image with graph filter banks, a graph is decomposed into two bipartite subgraphs (Fig. 2(b) and (c)). After filtering the graph signal along one bipartite subgraph with the graph filters, the results are stored in vertices, and filtering operation is applied to the resulting graph signals following the edges of the other bipartite subgraph. If we use critically sampled graph filter banks on the image signal, the diagonal edges will be ignored in a single stage if only the rectangular bipartite graph is used. Moreover, horizontal and vertical edges will be ignored if only the diagonal graph is used. But in the work [5], the oversampled graph Laplacian matrix can solve these problems. We append diagonal edges to the rectangular bipartite graph while keeping the oversampled graph bipartite as shown in Fig. 2(d). And then we use the oversampled diagonal bipartite graph (Fig. 2(e)) as a second stage of decomposition. Therefore, for oversampled bipartite images, we can use rectangular images with diagonal connection and diagonal graphs with horizontal and vertical connection to transform images in single stage transform.

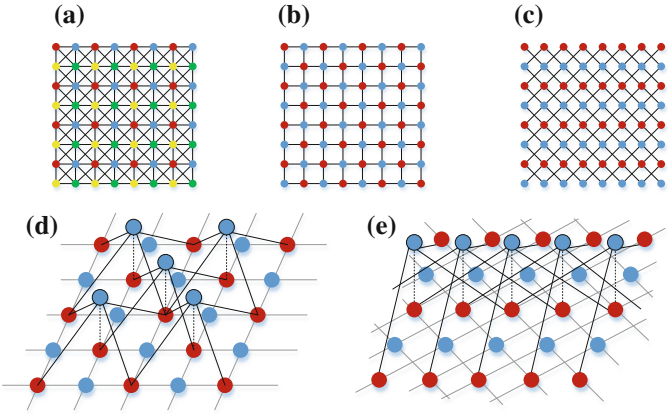


Fig. 2. (a) Image graph. (b) Rectangular bipartite subgraph. (c) Diagonal bipartite subgraph. (d) Oversampled rectangular bipartite graph. (e) Oversampled diagonal bipartite graph. The appended nodes are black circles filled with blue, and the appended edges are black lines.

In this paper, we consider an image as a signal on the regular graph. We use the oversampled graph filter bank to decompose NIR, LWIR, and visible images. Assume that I_1, I_2 and I_3 represent NIR, LWIR, and visible images, respectively. For the L-level decomposition, their low frequency subbands are denoted as $GW_{I_1}^L$, $GW_{I_2}^L$ and $GW_{I_3}^L$. The h-th high frequency subband at the level l are denoted as $GW_{I_1,h}^l$, $GW_{I_2,h}^l$ and $GW_{I_3,h}^l$.

For the low frequency subbands fusion rules, we take the low frequency coefficients of three source images as absolute values and normalized. And than we define the matching degree $MNGW_{I_i}^L(x,y)$ as that taking the interpolation of the maximum and minimum values of each low frequency coefficient. Finally, we determine the fusion operator as follows. The matching threshold is defined as T (0.5–1).

$$\begin{cases} GW_F^L(x,y) = \sum_{i=1}^{3} p_{I_i} GW_{I_i}^L(x,y) & \text{if } MNGW_{I_i}^L \geq T \\ GW_F^L(x,y) = \sum_{i=1}^{3} NGW_{I_i}^L(x,y) GW_{I_i}^L(x,y) & \text{if } MNGW_{I_i}^L < T \end{cases} \tag{4}$$

where if a low frequency component satisfies the following conditions, $\left|GW_{I_k}^L(x,y)\right| = \max_{i\in\{1,2,3\}} \{GW_{I_i}^L(x,y)\}$, then $p_{I_k} = 1$, the rest $p_{I_i} = 0$ $(i \neq k)$.

For the high frequency subbands fusion rule is similar to the low frequency subbands. The difference is that we use the local coefficient of variation (LCV) to define absolute values and normalized. LCV is defined as the ratio of the standard deviation to the mean, which is often used to measure the dispersion of the data.

The above fusion rule shows that it can autoselect the fusion method of weighted average or maximum absolute value independently. This can clearly preserve the

detailed information of the image significant signal, avoid the loss of information, reduce the noise and ensure the consistency of the fused image.

Finally, the final fusion image is obtained by applying the inverse transform on the fusion subbands GW_F^L and $GW_{F,h}^l$.

4 Experiments and Performance Comparisons

In this section, we conduct the image fusion experiments on four sets of infrared and visible images, "sandpath", "bench", "Kaptein_1123" and "soldiers_with_jeep". In order to simplify the oversampled graph filter banks decomposition, we resize the image so that its width and height are the power of 2. We set the decomposition level as 3. In the fusion rules, the matching threshold T is determined to 0.8. The other parameter is the window size for LCV, which is set to be 5×5 in all experiments. In order to evaluate the effect of fusion image subjectively, four kinds of fusion quality indicators are used to objectively evaluate the fusion performance of images, including standard deviation (SD) [8], average gradient (AG) [8], entropy (EN) [8], spatial frequency (SF) [8]. For demonstrating the effectiveness of the introduced method, we compare it with other image fusion methods based on wavelet transform (SWT) [9], multiresolution SVD (MSVD) [10], spectral graph wavelet transform (SGWT) [4], contourlet transform (CVT) [11]. These methods can be extended to multi-image fusion, and the parameters are in accordance with the requirements of the corresponding references.

From Fig. 3(a) and (c), we can see that the proposed method fully combines the key information in the three images, for example, the texture of the trees in the NIR image, the target people in the LWIR image, and the roads and fences in the visible image. As shown in Table 1, our proposed method has the best fusion quality indicators SD and EN. Although the measure values of AG and SF of the fusion method CVT are the highest, the target people of LWIR image and the road of visible image have not been effectively fused into the final result in proposed method.

As for the source images of "bench", it is observed that the target person from the source LWIR image is the most highlighted in fourth column. But the target chair from the source NIR image is not clear compared with other methods. From the Table 2, we can see that our proposed method has the highest measure values SD and EN, and the fusion method CVT has the best values of AG and SF. However, in the last column, the target person and the lake is not clearest compared with the proposed method. From these aspects, our method has better fusion effect.

From the Table 4, we can see that our proposed method has the highest measure values of all the metrics. And in the fourth column, it is observed that the targets (the people and the jeep) from the source LWIR image is the most highlighted. However, the trees from the source NIR image and the cloud from the source LWIR image of our fusion result is not clear compared with other method. But from above all the conclusions we can sure that our proposed method has the best fusion performance evaluation index, and has the better visual effects.

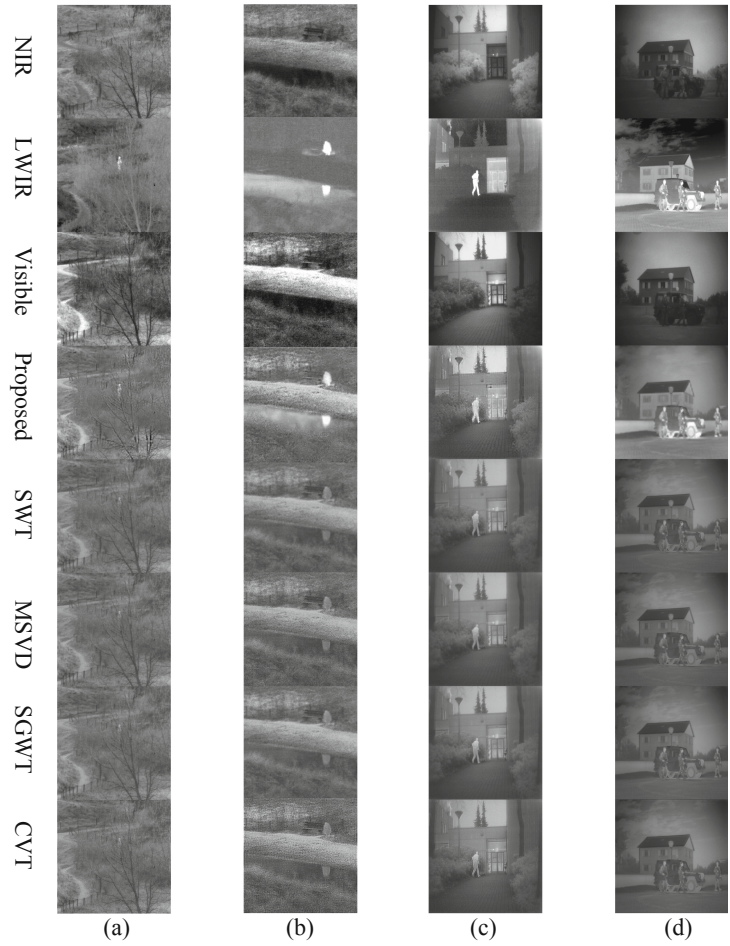


Fig. 3. (a)–(d) represent four sets of source images, “sandpath”, “bench”, “Kaptein_1123” and “soldiers_with_jeep”, respectively.

Table 1. Performance of different fusion methods on images “sandpath”

Metrics	Proposed method	SWT	MSVD	SGWT	CVT
SD	**21.4485**	16.2893	16.7374	16.6434	17.7302
AG	7.8655	5.7560	7.8326	5.4220	**9.0118**
EN	**6.2952**	6.0077	6.0503	6.0282	6.1421
SF	9.5459	6.7486	9.1424	6.2990	**10.3676**

By comparing the fusion results of the four sets of images, it can be seen that from the subjective effect, the proposed method can balance the intensity of NIR, LWIR and visible images, and retain the key information of the source images as much as possible.

Table 2. Performance of different fusion methods on images "bench"

Metrics	Proposed method	SWT	MSVD	SGWT	CVT
SD	**36.9120**	27.1114	28.9575	27.4300	30.0753
AG	14.5918	10.9863	16.6017	8.5558	**18.3572**
EN	**6.9553**	6.5196	6.6535	6.4974	6.7383
SF	18.3531	13.0310	20.3424	10.6604	**22.1066**

As can be seen from the Tables 1, 2, 3 and 4, the proposed method has the best SD and EN and relatively good AG and SF. This means that the fusion image with our method has higher contrast, sharpness and texture change, and receives more information from the source images. It can be obtained from the above conclusions that the proposed method in this paper has a very good visual effect and fusion quality.

Table 3. Performance of different fusion methods on images "Kaptein_1123"

Metrics	Proposed method	SWT	MSVD	SGWT	CVT
SD	**43.4076**	32.2459	32.2788	32.7235	32.8741
AG	**6.9620**	4.3349	4.6254	3.7465	6.7619
EN	**7.0422**	6.6608	6.6558	6.6668	6.7387
SF	**9.3150**	5.8321	6.2492	4.8416	8.4916

Table 4. Performance of different fusion methods on images "soldiers_with_jeep"

Metrics	Proposed method	SWT	MSVD	SGWT	CVT
SD	**31.8183**	19.2714	19.3691	19.4743	19.8146
AG	**4.1877**	2.5729	3.1174	2.1827	4.1584
EN	**6.9192**	6.2569	6.2674	6.2821	6.2979
SF	**5.6937**	3.3822	4.8030	3.0293	5.5961

5 Conclusion

In this paper, a new method based on oversampled graph filter banks for image fusion is proposed. First, we use oversampled graph filter banks operation on the bipartite graph to obtain low frequency and high frequency subband coefficients, which offset the effects of spectral folding. Then we proposed a new fusion rule to fuse these low subband coefficients and high subband coefficients, which can autoselect the fusion method of weighted average or maximum absolute value independently. Finally, the final fusion image is obtained by reconstructing the fused subband coefficients. This method not only retains most of the information of the source images, but also preserves the edge and texture information of the source image. We use four sets of NIR, LWIR, and visible images to verify the fusion performance of the proposed method and compare it with other four methods. The experimental results show that the new method proposed in this paper has better results in both visual analysis and objective evaluation.

Acknowledgments. The authors would like to thank the anonymous reviewers for their valuable revise opinion and suggestions in improving the technical presentation of this paper.

References

1. Lei, B., Yin, X.L., Li, X.Z.: Near infrared ray and visible light image fusion algorithm based on score. Comput. Eng. **39**(4), 226–229,233 (2013)
2. Xu, H., Wang, Y., Wu, Y., Qian, Y., Xu, H., et al.: Infrared and multi-type images fusion algorithm based on contrast pyramid transform. Infrared Phys. Technol. **78**, 133–146 (2016)
3. Zhan, L., Zhuang, Y., Huang, L.: Infrared and visible images fusion method based on discrete wavelet transform. J. Comput. **28**(2), 57–71 (2017)
4. Yan, X., et al.: Infrared and visible image fusion with spectral graph wavelet transform. JOSA **32**(9), 1643–1652 (2015)
5. Sakiyama, A., Tanaka, Y.: Oversampled graph Laplacian matrix for graph filter banks. IEEE Trans. Signal Process. **62**(24), 6425–6437 (2014)
6. Narang, S.K.: Critically sampled wavelet filterbanks on graphs. Dissertations & Theses – Gradworks (2012)
7. Narang, S.K., Ortega, A.: Perfect reconstruction two-channel wavelet filter-banks for graph structured data. IEEE Trans. Signal Process. **60**(6), 2786–2799 (2011)
8. Li, S., Kang, X., Hu, J.: Image fusion with guided filtering. IEEE Trans. Image Process. **22** (7), 2864–2875 (2013)
9. Rockinger, O.: Image sequence fusion using a shift-invariant wavelet transform. In: International Conference on Image Processing. IEEE Computer Society (1997)
10. Majumder, S.: Multiresolution SVD based image watermarking scheme using noise visibility function. Int. J. Appl. Evol. Comput. (IJAEC) **8**(1), 38–48 (2017)
11. Zhenfeng, S., Jun, L., Qimin, C.: Fusion of infrared and visible images based on focus measure operators in the curvelet domain. Appl. Opt. **51**(12), 1910–1921 (2012)

The Research of Regional Tourist Flow Situation Assessment Based on Time Variant and Multi-source Data

Yun Wu[1,2(✉)], Tian Fu[2], and Tian Ming Wang[3]

[1] College of Human Informatics and Cognitive Sciences, Waseda University, Tokorozawa, Japan
wuyunyunzhi@yahoo.co.jp
[2] Hainan College of Software Technology, Qionghai, Hainan, China
[3] Hainan College of Economics and Business, Haikou, Hainan, China

Abstract. At present, with the emerging of the independent travel, the tourist flow is equipped with stronger nonlinear feature. According to lots of researches, the multiple-source data integration could realize the higher tourist flow prediction accuracy than the prediction only based on the single source data. Targeting to regional passengers' multiple joints for travel space-time behavior as feature, the paper proposed the multiple-source data integration and explores to applies the situation awareness to the regional tourism flow prediction so as to formulate the neural network model based on the intelligent neuron component. Then, the author also utilizes this model to forecast the regional tourist flow as well as presses ahead the empirical researches by taking the tourist attractions of Hainan Province as example. By virtue of the experimental simulation, it analyzes the advantages of the prediction model than the prediction based on single source data.

Keywords: Search data · Situation Awareness · Multiple-source data · Neural network · Prediction model

1 Introduction

The tourism industry is one of the major resources of various countries' non-trade foreign exchange earnings and brings forth the direct and indirect contributions to the growth of their GDP (gross domestic product). According to the statistics of WTTC (World Travel and Tourism Council), in 2017, tourism industry's total contribution value reaches to 9.12 trillion RMB (1.3490 trillion dollars) to Chinese economy [1]. The travel flow prediction is the necessary and important link for the market, scheme and development process of tourism industry. Due to be affected by the climate, environment, holiday, emergent incidents and other related factors, the tourism is uncertainty. Particularly, with the independent travel emerging, tourism flow's nonlinear feature become clearer. Hence, the traditional prediction relying on the time sequence model and measurement model is difficult to forecast the tourist flow quickly and effectively. While the Internet application expands widely, the tourists could obtain various kinds of travel resources and information easily through the Internet.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 11–19, 2020.
https://doi.org/10.1007/978-981-32-9244-4_2

At present, network search engine and interactive platform are the major sources for tourists to gain the scene related information. As gaining more tourism information, the tourists would make the corresponding decisions. Thus, how to scientifically and effectively make the accurate and reliable situation prediction through multiple-source network related data not only can make up the deficiency of the lagging data sources of the tradition prediction method but also provide the management decision basis for scene management decision maker so that it plays a significant role in managing tourism resources and alleviating tourism environment pressure.

2 Research Status Analysis

The prediction based on network search data originated from an article about prediction which was issued on Nature by Ginsberg et al. [2]. With appearance of Google and Baidu search index as well as the wide utilization of Microblog, facebook, Twitter and other interactive platforms, lots of researchers apply the network data application into the studies of all industries. Korean scholar - Park S applied the Google trend search data to make the short-term prediction on the Japanese entry tourist data to Korea. He combined Google key words through mean absolute error as well as compared Autoregressive Moving Average (ARMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) on the basis of basic tourist flow data. From the researches, Google search data variation increases the prediction performance as well as improves the prediction accuracy [3]. In 2017, the scholar Li Xin et al. delivered a prediction article about Beijing monthly tourist flow combining Baidu search engine data [4]. Through the analysis on search and network data, Pan et al. [5] found the recent news about current tourism market timely and gained the better prediction results through remedying and supplying the traditional prediction model.

Up to now, most tourist flow prediction is direct to the single scene. From the integrative prediction research based on single network search data and traditional statistic data, they gain the obvious prediction accuracy from the long term prediction. However, with the independent travel rising, such models show some shortcomings for the effects of the multiple-jointed and short-term travel prediction. Hence, targeting to the features of the emerging family travel, self-driving tour, backpackers and so forth, this paper puts forward the tourism flow situation evaluation model based on time variant multiple data. On this basis, this approach assesses the scenic travel flow prediction under the dynamic variation as well as forecasts the travel route with the maximum probability so as to build up the perfect tourism security alarming system for the scene management decision makers. It plays a significant and critical role in realizing the tourism management, alleviating travel environment pressure and promoting the scientific development of local tourism.

3 Tourism Flow Prediction Model Based on Multiple-Source

For planning the multinode travel route, the tourists would comprehensively consider the scenic area evaluation, weather situation, consumption level, transportation condition and other factors. They always acquire the above-mentioned information through the search engine, Microblog and other platforms. The popular degree of the key words could directly and mostly reflect the tourists' attention to the scenic areas. So, analyzing the network key words could assist the tourists to understand the development trend of the travel flow.

3.1 Key Words Popularity Weight Ratio

Due to various personal preferences, the visitors could acquire the related scenic area information through the different platforms. Visitors' attention would directly show visitors' travel willingness. They possess certain distinctions on attention degree and timeliness. And then, the network search engines need to optimize the search results which cause the lag of the timeliness. But the Microblog is the visit experience during traveling which enjoy the stronger timeliness than the network search. Hence, the tourist would integrate the network search and Microblog interaction information to make the final decision. The paper aims to predict the scene tourism flow variation and to avoid the shortcomings that the single search data could not comprehensively and objectively key words' population. This paper launches the key word index expression method based on multiple sources so as to demonstrate the whole network trend of the key words. The formula is showed as following:

$$M = w_1 \cdot \left(\frac{|x_1 - u_1|}{\partial_1} \right) + \cdots + w_i \cdot \left(\frac{|x_i - u_i|}{\partial_i} \right)$$

In the formula, M is the dependable value of network key word popularity. On an occasion, for the key word popularity, the higher dependable value, the higher key word popularity. And the x_i is the key word index in some network search engine, u_i is the network index mean value, ∂_i is the standard deviation and w_i is network popularity weight ration during current period. The x_1 is key word's micro index, w_1 is network popularity weight ratio during current period, u_1 is the index mean value of Microblog key word and ∂_1 is the standard deviation. The weight value w is determined in line with the historical data statistics. Comparing with the single network search data, M dependable value could eliminate the one-sidedness influence of the single network platform as well as really reflect the whole network popular variation trend of the key words.

3.2 Key Words Empirical Research

3.2.1 The Selection of Key Words

Owning to more and more tourists acquiring the scenic area information through the Internet, accurate selection of the key words would affect the tourism prediction results directly, so the selection of the key words shall comprehensively reflect the overall

growth tendency of the research subjects. In order to analyzing visitors concerned tourism information contents, the paper utilizes the Baidu key words excavation tool to pick partial key words and then selects the key words which enjoy the similar fluctuation trend of the selected and prediction variables by clustering method to composite the key word index [6]. Finally, the thesis selects the 7 key words which are featured by the high repetition and high frequent research – "Bo'ao", "Wuzhizhou Island", "Yalong Bay", "Nanshan Temple", "Tianya Haijiao", "Clear Water Bay" and "The Weather of Hainan".

3.2.2 Data Verification

Data correlation analysis is to ensure whether two or more random variables with equal status to have the positive correlation or not. The collected data of the paper include the island inflow data and key word data and their value range is between Jan. 2014 to Dec. 2017. In addition, the island inflow data originates from the travel statistics published by Hainan Tourism Development Board. And the network key words statistics originates from Google trend website. 2014–2017 hainan the tourism flow data and key word popularity variation could be referred to Figs. 1 and 2. In order to ensure the close correlation between collected key words and tourist number, the paper utilizes SPSS software to calculate the correlation index between key words and scene tourist number. The results of the key word correlation test are showed as following Table 1.

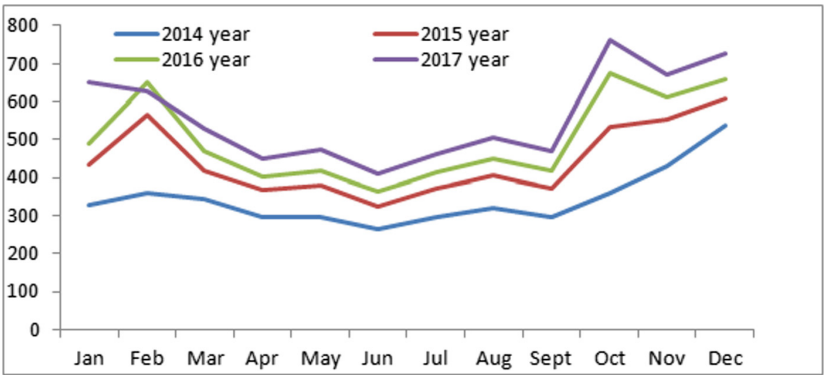


Fig. 1. Tourist flow linear

Granger causality test could verify a variable whether be favorable to explain another variable's prospect variation trend or not, namely judging whether the two variables are equipped with prediction capacity or not [7]. In order to select the independent variable with better prediction capacity, it needs to press ahead the Granger causality test on the selected independent variables. This paper sets the 7 key words - "Bo'ao", "Wuzhizhou Island", "Yalong Bay", "Nanshan Temple", "Tianya Haijiao", "Clear Water Bay" and "The Weather of Hainan" as X and the actual tourist number in scene as Y. And the Granger causality results between two variables are demonstrated in the following Table 2.

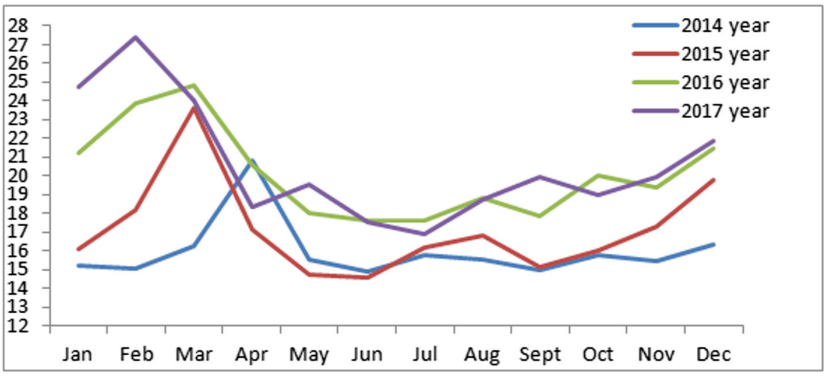


Fig. 2. Tourism key word popularity index linear

Table 1. The correlation index between key words and tourism flow.

Key word	Pearson correlation	Spearman correlation coefficient	p ratios	Conclusion
boao	0.671	0.613	0.00	Correlation
WuzhizhouIsland	0.738	0.726	0.00	Correlation
YaLongWan	0.751	0.763	0.00	Correlation
NanshanTemple	0.673	0.626	0.00	Correlation
Tianya Haijiao	0.759	0.637	0.00	Correlation
Clear Water Bay	0.801	0.784	0.00	Correlation
Hainan weather	0.7825	0.7175	0.00	Correlation

Table 2. Granger causality test results

Key word	F-Statistic	p ratios	Samples
Y does not Granger Cause X	0.8903	0.4271	48
X does not Granger Cause Y	4.52374	0.0168	48

By the above Table 3, the Granger causality test of the key words - “Bo’ao”, “Wuzhizhou Island”, “Yalong Bay”, “Nanshan Temple”, “Tianya Haijiao”, “Clear Water Bay” and “The Weather of Hainan” could state that the one-way causality relations between key word popularity and Hainan scene real tourist flow. It also shows that the key word index is armed with the prediction ability to tourist flow. Namely, if the 7 key words’ search indexes happen to variations, the statistics for Hainan tourism would change accordingly.

3.3 Situation Awareness Prediction Model Based on Generalized Neural Network

In 1999, for the first time, Tim Bass put forward the concept of Situation Awareness [8]. The core thought of Situation Awareness is to predict the incidents' prospect trends through collecting real-time related multisource data as well as relating with the relevant incident by analyzing related data so as to provide the data support for evaluating current management status.

This paper adopts the situation awareness into the tourism flow evaluation model, extracts tourism flow relevant perception factor information by analyzing tourism space-time flow feature to complete the understand about tourist space-time behaviors so as to realize the evaluation to the scene tourism flow situation. In line with the tourists' multimode tourism route analysis, the tourist would select a core scene as the major tourism node and formulate a multimode travel route integrating the surrounding scenic areas.

Through establishing high-order nonlinear prediction model, the neural network model's prediction accuracy is usually better than other traditional time sequence and measurement model [9]. Due to the scene node's trend flow state, key word popularity, scenic route flow, GPS path check popularity enjoying the positive dynamic change with the time, this paper utilizes the generalized neural network model (showing as followed Fig. 3) to make the analysis and evaluation on the variation of regional scene tourism flow status so as to reflect the scene flow evaluation to the decision manager to issue the corresponding measures for the coming flow peak, to optimize service process as well as improve the scene service quality.

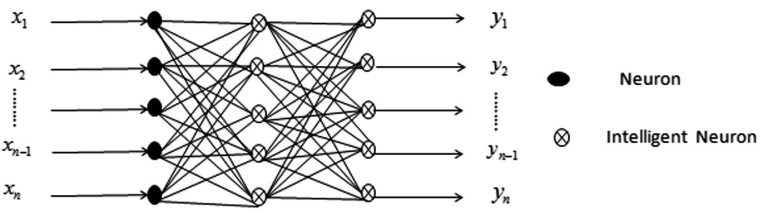


Fig. 3. The generalized neural network model

The intelligent neuron is the neuron to adjust the manipulation function by certain study rules as well as be equipped with the information storage. In the relevant researches, adopting linear independent function preprocesses the neural network input layer to make the neural network be better mapping effects [10]. The intelligent neuron contains lots of structures. And in this paper, the author uses the linear independent function to formulate the intelligent neuron [11]. The benefit of such intelligent neuron is that without increasing the variables in the neuron, it could add the dimension of variable by expanding function as well as enhance the neuron knowledge storage ability so as to gain better function mapping effects, referring to the following specific demonstration.

Inputting neuron as x to expand the linear independent function group to form the new neuron model, as shown in the Fig. 4.

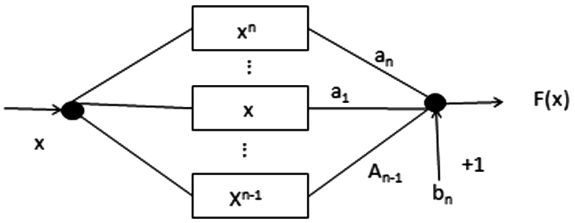


Fig. 4. Neuron model

The x is neuron input and a_r, b_n are the variable parameters. $f(x)$ is neuron output $f(x) = \sum_{r-1}^{n} a_r, x^r + b_n$. When the neuron has N samples, the input mapping formula is: $AX + b_n = Y$. In addition, the Y is the neuron output vector.

$$X = \begin{bmatrix} x_1, & x_1^2, & x_1^3, & \ldots, & x_1^n \\ x_2, & x_2^2, & x_2^3 & \ldots, & x_2^n \\ \vdots & & & & \\ x_n, & x_n^2, & x_n^3, & \ldots, & x_n^n \end{bmatrix}, \quad A^T = (a_1, a_2, \ldots a_n), \quad \text{Eliciting the following}$$

formula through the Vandermonde Determinant: $\prod_{r-1}^{n} x_r \prod (x_j - x_i),\ 1 \le i \le j \le n$.

During the model training process, adopting counterpropagation amending the weight and threshold value caters to error [12], as well as utilizing LMS algorithm regulates the adjustable parameters.

The stimulated training is applying the tourism flow relevant data: historical statistic data, network key word are the model training data which are used to test the precision of data and model prediction. And Then, the generalized neural network model composed of intelligent neurons is used to gradually predict the tourist flow state of the scenic spot in the T-time cycle.

4 Simulation Experiment Analysis

In the prediction process, the key words and historical statistical data of scenic spots are used to predict the tourist flow of the relevant scenic spots, and then Scenic spot prediction data and GPS data to carry out secondary training on the forecasting model, so that the tourist flow of the upstream scenic spots can be used to forecast the tourist flow of the specific downstream scenic spots. In addition, the island entry data come from the travel statistics published by Hainan Tourism Development Board. The island entry data are the month data and the data range are from Jan. 2014 to Dec. 2017. The search data come from the search trend data from various websites and then calculate the final key word statistic data in line with the above-mentioned key word popularity

weight ratio. The paper adopts the normalization processing to input samples. The six days' training data are selected from the "beginning", "middle" and "end" of each month, totally screening 288 sets data. Randomly selecting 260 sets data repeatedly trains the prediction data and the left 28 sets data are the verification model.

It predicts the variation of the adjacent scenic area tourist flow by the prediction model as well as forecasts the continuous 3 days scenic area tourist flow of Sanya, Hainan during National Day of 2018, as shown below (Table 3).

Table 3. The results of tourism flow prediction (ten thousands)

Date	Source Scenic				
	Actual flow	Predicted flow	Prediction error %	The target Scenic (一)	The target Scenic (二)
Ultima Thule 2018-10-2	3.28	3.06	6.71	Nanshan Temple	YaLongWan
Wuzhizhou Island 2018-10-3	0.74	0.71	4.05	Nanshan Temple	Luhuitou Park
Nanshan Temple 2018-10-4	4.27	3.25	22.3	Ultima Thule	Eternal love sanya

As shown in the Table 3, Keywords and statistical data were used to predict the scenic spot, with an error rate of about 7%, a maximum error of 8.02% and a minimum error of 4.05%. Because of the step-by-step prediction, the error of the downstream scenic spot is about 22%. With the popularization of Internet, if the categorical data enjoys the higher quality, the prediction values are closer to the true value. The neural network prediction model which uses the multisource and historical data enjoys a higher fitting degree and strong efficiency. Its prediction accuracy is higher than the historical traditional data prediction model, but the new prediction model also could improve the historical traditional data prediction model's such shortcoming greatly.

5 Conclusion

With the independent travel emerging and rising, due to be affected by weather, transportation, personal preference and many other elements, the regional scenic area tourist flow has a very strong non-linearity. The contributions of this paper are mainly establishing the regional scenic area tourist flow prediction model based on time variant and multisource as well as maximumly forecasting the tourism tourist flow variation route by relying on the change of the heat of GPS Path query change within the period. The experiments demonstrate that the model gains the better prediction results. Although the generalized neural network enjoys the strong adaptive learning ability and is able to stimulate the scenic area tourist flow regulation better, there are still many objective limitations for the real scenic areas. Hence, it needs further to make research on how to ensure the relevance among many involved factors of tourist flow to clearly and accurately predict the tourist flow.

Acknowledgments. During the visiting in Networked Information Systems Laboratory Waseda University. This work was supported by Hainan Natural Science Foundation of china under Grant No. 617172.

References

1. The tenth world tourism congress inventory [EB/OL], 1 July 2017
2. Ginsberg, J., Mohebbi, M.H., Patel, R.S.: Detecting influenza epidemics using search engine query data. Nature **457**(7232), 1012–1014 (2009)
3. Park, S., Lee, J., Song, W.: Short-term forecasting of Japanese tourist inflow to South Korea using Google trend data. J. Travel Tour. Mark. 1–12 (2016)
4. Lix, P.B., et al.: Forecasting tourism demand with composite search index. Tour. Manag. **59**, 57–66 (2017)
5. Pan, B., Wu, D.C., Song, H.: Forecasting hotel room demand using search engine data. J. Hosp. Tour. Technol. **3**(3), 196–210 (2012)
6. Zhang, L., Zhang, X., Cui, Y.: Research on keyword optimization and passenger flow prediction of Baidu search index based on clustering method. Manage. Rev. **8**, 126–134 (2008)
7. Tian, F., Zhen, W.: Scenic spot tourists flow prediction research based on web search items. In: 2nd Joint International Information Technology, Mechanical and Electronic Engineering Conference (2017)
8. Bass, T., Gruber, D.: A glimpse into the future of id, special issue intrusion detection. USENIX Assoc. Mag. (2005)
9. Kon, S.C., Turner, L.W.: Neural network forecasting of tourism demand. Tour. Econ. **11**(3), 301–328 (2005)
10. Eck, J.T., Shin, F.Y.: An automatic text-free speaker recognition system based of enhanced ART 2 neural architecture. Inf. Sci. **76**, 233–253 (1994)
11. Li, Y., Ding, Y., Wang, D.E., et al.: Research on the design methods of tourist routes in scenic spots with time constraints and spatial behavior characteristics of tourists. Travel J. **31** (9), 50–60 (2016)
12. Law, R.: Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. Tour. Manag. **21**(4), 331–340 (2000)

The Chinese Knowledge Graph on Domain-Tourism

Weizhen Zhang, Han Cao(✉), Fei Hao, Lu Yang, Muhib Ahmad, and Yifei Li

School of Computer Science, Shaanxi Normal University, Xi'an 710119, Shaanxi, China
{zhangweizhen, caohan, fhao, lu-yang, liyifei}@snnu.edu.cn

Abstract. Tourism plays an increasingly important role in people's daily life. However, in the era of big data, tourists find, it is difficult to acquire useful knowledge of travel information on the Internet. Knowledge Graph describes the real-world concepts, entities and their relationships in a structured form. Which is capable of mapping the Internet information in the form that is closer to the human cognition. Thus, provides the ability to organize, manage and understand the vast travel information available on the Internet. In order to promote the sharing of tourism knowledge and culture in China, we are committed to constructing the Chinese knowledge graph on domain-tourism, in which knowledge is obtained from the existing Chinese encyclopedia knowledge graph and unstructured web pages. In addition, we provide a storage scheme of RDF triples data in the graph database and build the tourism knowledge application platform based on it.

Keywords: Knowledge graph · Tourism · RDF

1 Introduction

Recently, with the development of Semantic Web, knowledge representation has evolved from early first-order logic (e.g., FOIL [1], which learns probabilistic Horn clauses.), production rule to RDF and OWL. Based on these representations, a large amount of structured knowledge in different fields is generated and published by building a common ontology base [2]. Knowledge graph based on semantic web, which is further divided into two categories: general-purpose knowledge graph (GKG) and domain-specific knowledge graph (DKG). General-purpose includes WordNet [3], DBPedia, YAGO [2], Freebase, etc. The study of Chinese knowledge graph can be traced back to HowNet project in China. In the industry, there are OpenKG.CN, Baidu Zhixin, Sogou Zhicube. The academic circle includes XLore, zhishi.me [4], and CN-DBpedia [5]. However, there are few Chinese domain-specific knowledge graph. Especially, there is still shortage of Chinese knowledge graph on tourism domain, which seriously hinders the development and inheritance of Chinese

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 20–27, 2020.
https://doi.org/10.1007/978-981-32-9244-4_3

tourism culture. Therefore, we are committed to build a Chinese domain specific tourism knowledge graph. The key contributions of this work are concluded as follows:

1. We have studied the neural network for word vector representation model in natural language processing (NLP), and these models use to generate Chinese word vectors to achieve the purpose of entity alignment.
2. After acquiring tourism knowledge, The Chinese domain-tourism ontology is constructed by protégé and save it as RDF triples.
3. We map the tourism knowledge to the ontology, then the RDF triples data built and stored in the Neo4j graph database. And finally based on the graph database a tourism knowledge application platform built.

The rest of this paper is organized as follows. In Sect. 2, the construction of Chinese domain-tourism knowledge graph is described. Section 3 describes the construction of the domain-tourism ontology, the storage of tourism ontology and the knowledge application platform developed on it. Finally, we conclude the paper and point out further work in Sect. 4.

2 Construction of Chinese Domain-Tourism Knowledge Graph

Figure 1 shows the development architecture of our Chinese domain-tourism knowledge graph. It mainly includes three parts: acquisition of tourism knowledge, knowledge fusion and knowledge completion. And we develop the Chinese domain-tourism knowledge application platform based on it.

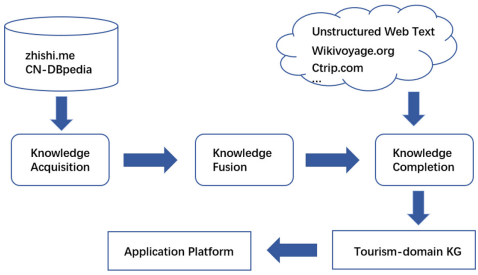


Fig. 1. Development architecture of Chinese domain-tourism Knowledge graph.

2.1 Knowledge Acquisition

The key data sources of this work are as follows: 1. Zhishi.me[1]: Zhishi.me provides dump download. This study uses turtle format data (RDF triple format). We obtain the links of tourism-related entities by searching for keywords such as travel, attractions,

[1] http://openkg.cn/dataset/zhishi-me-dump.

and scenic spots from the official website of zhishi.me. Then we extract the relevant knowledge from the downloaded data. The main knowledge extracted includes: abstract, category, infobox property and label. 2. CN-DBpedia is a large-scale general-purpose structured knowledge base. Its official website provides free data download service. We extract the travel-related knowledge from the structured triplet data that we downloaded.

2.2 Knowledge Fusion

Through knowledge acquisition, we obtain the entity, relationship and entity property information from data. After obtaining these goals, we need to integrate them to combine the properties, texts, and relationship information refer to same entity in different knowledge bases. The work we have done in this step focuses on entity alignment in knowledge fusion. After the entity alignment, we fuse the relevant information of the same entity and finally save it as triple.

Entity Alignment. Entity alignment refers to the process of judging whether there are other entities in the knowledge base with the same meaning, that is, whether one or more entities in the knowledge base points to the same object in the real language context [6]. The entity in the knowledge base may have multiple expressions, such as: different entity names of "terracotta army" and "terracotta warrior", and these different entities map to the same entity through entity alignment.

We have done vast study on word vector representation model in NLP, by using these models to calculate the cosine distance among entities in the semantic space, that is, the semantic similarity to get the purpose of entity alignment.

Skip-Gram Model. The Skip-gram model predicts the context through the target word, and the goal is to find word representation that is useful for predicting surrounding words in a sentence or a document [7]. More detail, given a sequence of training words: $w_1, w_2, w_3, \ldots, w_T$, the purpose of the Skip-gram model is to maximize the average log probability

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{0<|j|\le c}\log p(w_{t+j}|w_t) \tag{1}$$

where c denotes the size of the training window. The primary Skip-gram defines $p(w_{t+j}|w_t)$ using the softmax function:

$$p(w_0|w_1) = \frac{exp(v'^{T}_{wo}\, v_{w_I})}{\sum_{w-1}^{W} exp(v'^{T}_{w}\, v_{w_I})} \tag{2}$$

where v_w and v'_w represent the input and output vector representations of w respectively, and W is the total number of words. Mikolov also proposes using Hierarchical softmax or negative sampling category to improve the Skip-gram model [8].

Dataset. The web pages in Sogou-T[2] and Chinese Wikipedia dump[3] as the text corpus are used. Sogou-T is provided by a Chinese commercial search engine, which contains 2.7 billion words in total and Chinese Wikipedia dump contains approximately 2 billion words.

Settings. During training, the dimensions of word are dynamically selected ranging from 100 to 500 randomly. And we set the training iterations to be 100, 150 or 200. The word frequency less than 5 in the corpus is discarded, the Hierarchical Softmax algorithm is used as the improvement of the Skip-gram model, and for learning rate η, its initial value is 0.025.

Word2vec[4] is an efficient toolkit of Google to obtain word vectors that released in 2013. The implementation of the Skip-gram model in this paper is based on the toolkit.

In this paper, total 7 groups of experiments are carried out, and we select 95 tourism-related entities as sample words. The word vector (200 dimensions) of 61 entities is obtained through the training experiment with 200 training iterations. By calculating and obtaining the most similar entities, 36 entities aligned, with an accuracy rate of 59%. To exemplify the learned representations, in below we show the most similar entities for some sample entities in below Table 1. It proves that we have obtained the same pointed entity in the real environment, that is, the effect of entity alignment. For instance, "terracotta army" and "qin terracotta army", "terracotta warrior" all points to the same entity in real life, similarity "temple" and "monastery", "shrine" referring to the same entity. The CBOW model is similar to the SG model, so we are not repeating it again.

Table 1. Word Similarity result when trained using the Skip-gram model.

Entity	兵马俑(terracotta army)	度假(vocation)	旅游(tourism)	寺庙(temple)
Top Similar Entities	秦兵马俑(qin terracotta army)	渡假(holiday)	观光旅游(sightseeing tour)	寺院(monastery)
	陶俑(terracotta figurine)	度假胜地(vacation spot)	旅游业(tourism industry)	佛寺(buddhist temple)
	秦俑(terracotta warrior)	度假村(resort)	特色旅游(characteristic tourism)	庙宇(shrine)
	俑(tomb figure)	度假地(holiday resort)	观光(sightseeing)	该寺(the temple)
	秦始皇陵(qin shi huang mausoleum)	休闲(leisure)	旅游观光(tour and sightseeing)	道观(taoist temple)
	俑坑(figurine pit)	旅游胜地(tourist attraction)	生态旅游(ecotourism)	宫观(taoist shrine)
	秦始皇(qin shi huang)	避暑(prevent sunstroke)	旅游景点(tourism spot)	该庙(the shrine)
	铜车马(bronze chariot and horse)	旅游(tourism)	自驾游(self-driving tour)	庙(temple)
	汉墓(hantomb)	胜地(famous spot)	旅游区(tourist area)	寺观(taoist temple)
	古墓(ancient tomb)	疗养(recuperation)	红色旅游(red tourism)	清真寺(mosque)

Directional Skip-Gram Model. The DSG model of Tencent AI Lab is based on word vector training model of SG. On the basis of the co-occurrence relationship of words in the text window, an additional position vector is added to represent the relative position

[2] https://www.sogou.com/labs/resource/t.php.

[3] https://dumps.wikimedia.org/zhwiki/.

[4] https://code.google.com/p/word2vec/.

of target words in the given context, to improve the accuracy of semantic representation of word vectors. DSG model proposes an additional directional softmax function:

$$g(w_{t+j}, w_t) = \frac{exp\left(\delta_{w_{t+j}}^T v_{w_t}\right)}{\sum_{w=1}^{W} exp\left(\delta_{w_{t+j}}^T v_{w_t}\right)} \tag{3}$$

To measure the context w_{t+j} is more biased to the left or right of the target word w_t, where W denotes the total number of words in the corpus. v_w represents the word vector, and δ represents the direction vector of each word in the context [9]. Moreover, the g function has similar update strategy to the negative sampling technique: increase the probability of positive samples while reduce the probability of negative samples.

The final softmax function of DSG model is added to formula (2): $f(w_{t+j}, w_t) = p(w_{t+j}|w_t) + g(w_{t+j}, w_t)$.

We use the open source word vector published by the Tencent AI Lab. Using Chinese Wikipedia redirection (synonym) entity pairs as test data, total 22797 pairs, by computing the cosine distance of entity pairs to characterize the result of the entity alignment. In our experiment, the accuracy of entity alignment results is obtained by setting different thresholds of similarity. From the experimental results, when the threshold of similarity is set to 0.65, the accuracy rate of entity alignment as high as 91.67%. However, as the threshold continues to increase, the accuracy rate of entity alignment reduced, and when the threshold is set to 0.85, the accuracy rate becomes reduced to 45.26%.

BERT Model BERT addresses the Generative Pre-trained Transformer unidirectional constraints by proposing a new pre-training target. Model named "masked language model" (MLM) [10]. The masked language model masks randomly some of the tokens from the input, and predict the original vocabulary id of the masked word based only on its context [11]. Furthermore, BERT adds an additional sentence-level continuous prediction task on the basis of bidirectional language model, aiming at predicting whether the text at both ends of the input is continuous text. As a result, BERT representations are fine-tuned with just one additional output layer to create state-of-the-art models for a variety of tasks, without substantial task-specific architecture modifications [11].

We achieve entity alignment by using Google's pre-training BERT (BERT-Base, Chinese) model based on character level to get the word vector of entities. 22,797 pairs of Chinese Wikipedia entity pairs are extracted by us and HIT IR-Lab Tongyici Cilin (Extended) dataset, that are used for experimental data. When the thresholds are set to 0.65, 0.70, 0.75, 0.80 and 0.85, respectively, the accuracy of entity alignment is more than 99%. We also conduct a comparative experiment with the three models mentioned above. For a fair comparison among different models, we use same dimension size for all word embeddings, and discard the rare words that appeared less than 5 times in the training corpus. For learning rate η, its initial value is 0.025. The window size and negative samples are both set to 5. The experimental results are shown in Table 2.

Table 2. Accuracy of using different models in entity alignment. CWEP represents Chinese Wikipedia entity pairs dataset and HTTTCL represents HIT IR-Lab Tongyici Cilin (Extended) dataset.

	SG		CBOW		DSG		BERT	
0.65	0.58583	0.808177	0.522655	0.741696	0.902535	0.916743	0.99698	0.998947
0.7	0.482811	0.727069	0.439567	0.664587	0.808676	0.85599	0.99547	0.997894
0.75	0.369222	0.623766	0.347652	0.566432	0.767733	0.767733	0.997456	0.997456
0.8	0.259698	0.500329	0.246911	0.446793	0.690393	0.634031	0.994425	0.997061
0.85	0.151336	0.355442	0.150288	0.301582	0.358107	0.452603	0.987805	0.996008
	CWEP	HITTCL	CWEP	HITTCL	CWEP	HITTCL	CWEP	HITTCL

From the experimental results, compared with SG and CBOW models, DSG model adds an additional position vector to represent the relative position of target words in the given context, which improved the accuracy of semantic representation of word vectors. BERT model through pre-train deep bidirectional representations by jointly conditioning on both context in all layers, obtains state-of-art results on accuracy of semantic representation of the word vector. All these ideas have reference significance for us to study entity alignment.

2.3 Knowledge Completion

In the process of knowledge completion, the work we have done is mainly to complete the infobox properties of entities. Infobox properties is a form of knowledge, which is used to summarize the features of entities [12]. As the properties extracted from zhishi.me and CN-DBpedia are limited, we extract the properties related to tourism entities from the unstructured network text to complete the infobox properies. We use regular expressions in Python scripts to extract entity properties information from tourism-related sites (e.g. wikivoyage), mainly by making rules manually.

3 Construction of Chinese Domain-Tourism Ontology and Application Platform

To acquire structured knowledge, it is necessary to construct the ontology. By summarizing the concepts, properties and relations in the data, we determine the category structure of the domain-tourism ontology, combined with the ontology construction method of Stanford university, “seven-step method”, and use the ontology editing tool Protégé to complete our Chinese domain-tourism ontology construction.

Furthermore, the ontology saves in RDF format. RDF data usually appears in form of triples (S, P, O), i.e., (subject, predicate, object). An entity is usually described by multiple triplet information, so an entity triplet information can form an RDF directed subgraph, each triple is represented in the graph as “node-edge-node” relationship. Neo4j is a graph-oriented database whose primitives are nodes, relationships, and attributes. In this paper, we create a mapping between the labels of RDF data and the Neo4j database, store the RDF data in the Neo4j graph database.

Finally, we developed a domain-tourism knowledge application platform based on knowledge graph. The main functions are as follows: (1) Visualize the urban tourism heat and color depth represents the heat of tourism; (2) Semantic search: combine the knowledge base to parse the entities entered by users, so that users' intentions can be more accurately understood, and entity overview can be given in the form of knowledge cards; (3) Providing tourism-domain knowledge service, in which the specific scenic spots will display the deep knowledge information such as history, culture, geography, climate, sightseeing route, food, accommodation, etc.

4 Conclusion and Future Work

In this work, after information extraction, we introduce neural network for representation models to get the word vectors of the tourism entities, which are used to achieve the purpose of entity alignment. Combined with other knowledge graph technology we build the domain-tourism knowledge graph and develop a knowledge application platform.

This paper is beneficial exploration and practice from tourism information service to tourism knowledge service. In the future, in-depth research will be conducted on the aspects of question answering, accurate tourism entity links and automatic construction of tourism ontology, and recommendation systems.

Acknowledgements. This work is supported by the primary research and development plan of Shaanxi Province "Research on Characteristic Module Design of SOTA Platform for Intelligent Tourism" (NO: 2018SF-361), the National Key R&D Program of China under grant No. 2017YFB1402102, and the Fundamental Research Funds for the Central Universities (GK201806012).

References

1. Quinlan, J.R., Cameron-Jones, R.M.: Foil: a midterm report. In: Proceedings of ECML (1993)
2. Suchanek, F.M., Kasneci, G., Weikum, G.: Yago: a core of semantic knowledge. In: Proceedings of the 16th International Conference on World Wide Web, pp. 697–706. ACM (2007)
3. Fellbaum, C. (ed.): WordNet: An Electronic Lexical Database. MIT Press, Cambridge (1998)
4. Niu, X., Sun, X., Wang, H., Rong, S., Qi, G., Yu, Y.: Zhishi.me: weaving chinese linking open data. In: International Conference on the Semantic Web. Springer, Heidelberg (2011)
5. Xu, B., Xu, Y., Liang, J., Xie, C., Liang, B., Cui, W., et al.: CN-DBpedia: a never-ending chinese knowledge extraction system. In: International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems. Springer, Cham (2017)
6. Hao, Y., Zhang, Y., He, S., Liu, K., Zhao, J.: A joint embedding method for entity alignment of knowledge bases. In: Knowledge Graph and Semantic Computing: Semantic, Knowledge, and Linked Big Data. Springer, Singapore (2016)

7. Collobert, R., Weston, J.: A unified architecture for natural language processing: deep neural networks with multitask learning. In: International Conference on Machine Learning (2008)
8. Mikolov, T., Sutskever, I., Chen, K., et al.: Distributed representations of words and phrases and their compositionality. Adv. Neural. Inf. Process. Syst. **26**, 3111–3119 (2013)
9. Song, Y., Shi, S., Li, J., Zhang, H.: Directional skip-gram: explicitly distinguishing left and right context for word embeddings. In: NAACL 2018 (Short Paper)
10. Taylor, W.L.: Cloze procedure: a new tool for measuring readability. Journal. Bull. **30**(4), 415–433 (1953)
11. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: BERT: pre-training of deep bi-directional transformers for language understanding (2018)
12. Wu, T., Gao, C., Qi, G., et al.: KG-Buddhism: the Chinese knowledge graph on Buddhism (2017)

An Improved Change Detection Based on PCA and FCM Clustering for Earthen Ruins

Changqing Wang[1], Yun Xiao[1(✉)], Baoying Liu[1], Dexin Du[2], and Rui Luo[1]

[1] School of Information Science and Technology, Northwest University, Xi'an, China
{wchangqing, rui_ui_l}@stumail.nwu.edu.cn, {yxiao, paola.liu}@nwu.edu.cn
[2] Hanguangmen Site Museum of Xi'an Tang Imperial City Wall, Xi'an, China
252343654@qq.com

Abstract. Nondestructively detecting the damage change is important for the protection of the earthen ruins. How to effectively detect the subtle changes in earthen ruins is an urgent problem to be solved. In our paper, we use the difference and the log-likelihood ratio method to create a difference image, which can effectively avoid the influence of noise. Then the orthonormal eigenvectors are extracted through principal component analysis (PCA) of non-overlapping block set to create an eigenvector space which is mapped to each vector in turn to form a feature vector space. The feature vector space is partitioned into two clusters according to the feature vector approximate degree by using fuzzy c-means (FCM) clustering. The experiment results on the images of Hanguangmen earthen ruin show that this method can find the changing area simply and efficiently.

Keywords: Change detection · PCA · FCM clustering · Earthen ruins

1 Introduction

As a typical earthen ruin, the Tang Imperial City Wall is an important part of Chinese civilization. In order to protect the earthen ruin, it is important to assess the damage severity of the earthen ruin. So how to nondestructively detect the damage of earthen ruin is a big challenge in earthen ruins protection. The use of images to detect changes in earthen ruins provides us with an effective way [1]. Thus how to detect damage change accurately and effectively becomes an urgent problem to be solved.

The change detection technique identifies regional variations by comparing image changes at different times but in the same region. For different detection target objects, the current detection method can be divided into Pixel-Based Change Detection (PBCD), Object-Based Change Detection(OBCD) and mixed change detection Method (Hybrid Change Detection, HCD) [2–4]. PBCD generates a difference map with a single pixel as a basic processing unit and classifies the difference map. They are greatly affected by the inherent speckle noise of the image, and the detection accuracy is limited [5].

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 28–35, 2020.
https://doi.org/10.1007/978-981-32-9244-4_4

Turgay Celik proposes a method for unsupervised change detection images change using principal component analysis(PCA) and k-means clustering [6]. Orthonormal eigenvectors are extracted through PCA non-overlapping block set to create an eigenvector space. The change detection is achieved by partitioning the feature vector space into two clusters using k-means clustering. However, the difference image is easy to be affected by noise and the robustness of change detection is limited.

Leng et al. propose an improved hybrid change detection method to improve detection accuracy [7]. The method firstly uses the pixel-level change detection method to extract the initial change region. The difference map is generated by the average likelihood ratio [8]; then the fuzzy clustering (FCM) is used to change The SAR images are divided into three categories.

MAY Stéphane proposes an original method for transforming transform detection problems into multi-temporal image sequences [9]. The original sketch of the image is implemented using the improved Perona-Malik filter algorithm [10]. This method can effectively remove the edge noise information for the change results. Then, an application for change detection is proposed. Damage changes such as weathering in earthen ruins are slow, which often lead to subtle changes in two images of the same object at different times. So for the image of earthen ruins, the above methods are not suitable for detecting subtle change.

we propose an improved change detection method based on PCA and FCM clustering algorithm. Using the weighted difference image constructed with multi-temporal images, our method can automatically analyze the change of difference image. The difference image combines difference and the log-likelihood ratio value of the two images. Eigenvectors of pixel blocks from the difference image are extracted by PCA. Subsequently, a feature vector is constructed for each pixel in the difference image by projecting the pixel's neighborhood onto the eigenvectors. Then, the image can be divided into unchanging regions and different changing regions by FCM clustering. In order to avoid the edge effect, we also perform custom operator filtering on the classification results to obtain the final change detection results. Finally, we verify the effectiveness of our method through experiments.

The main contributions of this paper are:

(1) An improved image change detection method is proposed. We use PCA to extract the weighted difference image main feature vector to form the vector space. The FCM method is used to classify the vector space to form the change image.
(2) The proposed method can be effectively applied to most images with weakly pixel changes and it has strong anti-noise performance.

2 Change Detection Based on PCA and FCM Clustering

Our method is mainly divided into three parts. Firstly, the difference image and feature vector are obtained. The next step is to construct the spatial feature vector. Finally, the change image is obtained by clustering. The area of change is calculated according to the size of the region.

2.1 Weighted Difference Image

In general, the difference image has the absolute-valued differences of the intensity values of the corresponding pixels of the grayscale images, $X_{d1} = X_1 - X_2$, X_1 represents the initial reference image, X_2 represents the changed image.

However, the difference image obtained by the direct difference method has poor robustness. The ratio calculation is more suitable for the statistical characteristics [11]. Therefore, we use the local average logarithmic likelihood ratio change detection operator to generate the difference image X_{d2} and use the small blocks to replace the local homogenous region [12]. The simplified calculation formula of the average logarithmic likelihood ratio can be expressed as:

$$x_{d2}(i,j) = K \ln \frac{\sum\limits_{(m,n)\in V_{i,j}} X_2(m,n)}{\sum\limits_{(m,n)\in V_{i,j}} X_1(m,n)}, \tag{1}$$

where V represents the local neighborhood pixel set of the pixel (i,j) within the $W_1 \times W_1$ blocks(W_1 typically takes a value of 5). K is the amplification factor, and map image pixel values to [0, 255].

In order to ensure the accuracy of the difference and to avoid the influence of noise, we carry out the weight ratio fusion image according to the degree of image change.

$$X_d = \omega X_{d1} + (1-\omega) X_{d2}, \tag{2}$$

where ω represent the weight of X_{d1}, $0 \le \omega \le 1$. It has been found through testing that the greater the degree of image change, the better the direct difference effect, and vise visa.

2.2 Building the Feature Vector Space (FVS)

The next step is partitioning into $W_2 \times W_2$ non-overlapping blocks, where $W_2 \ge 2$, generally take odd numbers. Consequently, let $X_d(i,j) = \{x_d(m,n) | (m,n) \in V_{i,j}\}$ be $W_2 \times W_2$ difference image block referenced by coordinates (i,j). Notice that (i,j) is not continuous, and our default image can be divided equally by the block. Using the ordering of row-by-row extraction elements, the row vector formulation $x_d(i,j)$ can be obtained. For a $M \times N$ matrix, we can convert it into a matrix of $R = (M \times N)/(W_2 \times W_2)$ rows and $C = W_2 \times W_2$ columns. The average vector of the set is defined by

$$\phi = \frac{1}{R} \sum_{i=1}^{R} x_d^i \tag{3}$$

Each vector differs from the average vector by the vector $\Delta_i = x_d^i - \phi$. We use PCA to find a set of N orthogonal normalized vectors and their associated scalars. Orthogonal vector is defined as:

$$N = \frac{1}{R}\sum_{i=1}^{R} \Delta_i^T \Delta_i \tag{4}$$

The vector ξ_i and λ the corresponding scalars are the eigenvectors and eigenvalues of the covariance matrix, respectively. We arrange their eigenvalues in descending order, take the eigenvector $\xi_{i\max}$ corresponding to its largest eigenvalue. Finally, projecting the $\xi_{i\max}$ of each row vector $x_d(i,j)$ to construct feature vector space F_d. Every row vector in F_d be calculated by $f_i = \xi_{i\max}^T \Delta_i$, $1 \le i \le R$. The new feature vector space is defined as:

$$F_d = [f_1, f_2, \ldots, f_R]^T \tag{5}$$

2.3 FCM Clustering and Change Image

The feature vector contains important information about pixel change. We need to group the pixels into two disjoint classes. When dealing with the data stream, inheriting the approximate degree matrix can make traditional FCM more effective for data stream [13].

For FCM clustering, the mean value of the difference image intensity is taken as the first clustering center C_1. The shortest distance between each pixel point and the nearest cluster center is calculated. The new pixel point is selected as the new center according

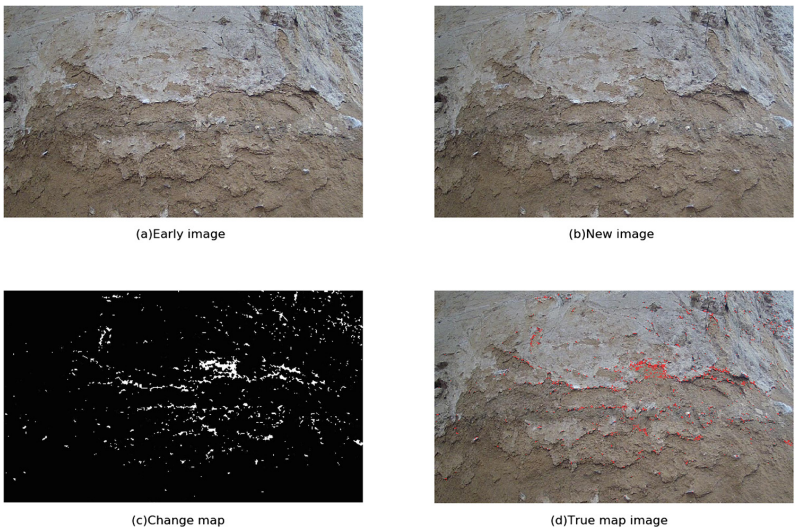


Fig. 1. The original image data and the change detection result are shown in the figure. (a) is the original image in September 2017, (b) is the new image in November 2018, (c) change result image with our method, (d) is the true map image, and red pixels represent the change parts in actual wall area.

to the probability distribution $dis(i,j)^2 / \sum_{x \in X} dis(i,j)^2$ until cluster center points C_2 are selected. The approximate degree of every pixel can be calculated by:

$$\varphi_{ij}(k) = \frac{1}{\left[d^2(x_{ij}, c_1)/d^2(x_{ij}, c_1) + d^2(x_{ij}, c_1)/d^2(x_{ij}, c_1)\right]^{1/(m-1)}}, \tag{6}$$

where φ_{ij} represents the approximate of the pixel sample x_{ij} to C_1, m stands for ambiguity, and d is Euclidean distance. The fuzzy index m takes a value of 2, the tolerance error $\varepsilon = 0.0001$.

After the iteration, the data is assigned to different classes according to an approximate degree. In the difference image, the cluster with lower average value is assigned as the unchanged class and another cluster is assigned as change class. We can build a binary image to show the result of the change detection. In order to eliminate the edge effects during image processing, we perform edge etching on the output image. The structural elements used for the etching operation use a 5×5 square structure element that a center diamond area value is 1 and a corner edge area is 0. Finally we output our results of the change results.

3 Experimental Results

In order to nondestructively detect the damage of earthen ruins, we have monitored the wall surface of Hanguangmen earthen ruin of the Tang Imperial City Wall with cameras twice a day since September 2017. We get 700 images of Hanguangmen earthen ruin at different times while in the same positions. In our experiments, we took two pictures before and after the change in two years. The typical parameters of block size is 5×5, different weight $\omega = 0.8$. We compare it with the method in [6] and evaluate it from the effects of change and robustness.

We describe the effects of our algorithm from an analytical and quantitative perspective and verify the effectiveness of the proposed method with these images.

Figure 1(a) and (b) are the reference image and the change image (size 1200×650) in the same area of Hanguangmen earthen ruin in September 2017 and November 2018 respectively. There are some subtle changes in the two images at different periods due to the destruction of the natural environment. Figure 1(c) is the final image change detection result. White parts indicate the change area, we can observe the change of the earthen ruin. Figure 1(d) is the change map image, and red parts represent the change area of Hanguangmen earthen ruin. By observing the results of the experiment, we can observe that the changing area is located at the edge of the damaged area. This follows the general rule of destruction of earth ruins. We can combine disease environment information with the degree of change in each region to regulate the environment and protect the severe areas.

We also measure the size of the change area according to the image size. The proportional relationship shows that the size corresponding to each pixel is about 0.07 cm^2. Therefore, according to the number of changed pixels, the area of the change area can be estimated to be 678.72 cm^2.

In order to verify the robustness, we use the original algorithm as a comparative quantitative analysis of the ability to resist noise. We use the peak signal-to-noise ratio (PSNR) to detect noise immunity. It is often simply defined by mean square error (MSE). If two $M \times N$ gray images I and K of noise approximation, then their mean square error is defined as:

$$PSNR = 10 \cdot \log_{10}(\frac{A^2}{MSE}), MSE = \frac{1}{MN}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}||I(i,j) - K(i,j)||^2. \tag{7}$$

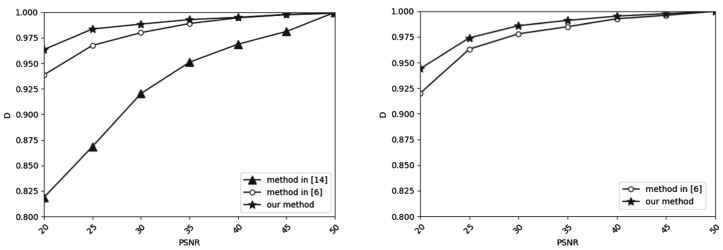
where A is the image maximum intensity. We set it to 1.

We use the original image X_1 and the changed image X_2 to generate a change image CI_1. Then we add noise generation to the X_1 image and generate a change image CI_2 with X_2. The degree of difference between CI_1 and CI_2 is defined as:

$$\& = 1 - \frac{1}{A^2}\frac{1}{MN}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(CI_1(i,j)||CI_2(i,j)). \tag{8}$$

where $(CI_1||CI_2)$ stands for "and" operates. If $CI_1 = CI_2$, $CI_1||CI_2 = 1$. Otherwise, $CI_1||CI_2 = 0$.

As shown in Fig. 2, it is a curve of algorithm comparison. we perform an anti-noise contrast experiment on the proposed algorithm and other algorithms. In Fig. 2(a), we compare our method with the methods [6, 14] by using zero-mean Gaussian noise. In Fig. 2(b), we make a speckle noise comparison with the [6] method. It can be seen that our method has better performance than that in [6, 14] when the PSNR value becomes large. It can be observed that the experiment of our improved algorithm is significantly better than the original algorithm and can better adapt to the images of earthen ruins.



(a) 0-means Gaussian noise (b) speckle noise

Fig. 2. Anti-noise experiments result are performed on the images, D represent similarity after adding noise. (a) is a 0-means Gaussian noise comparison, and (b) is a speckle noise comparison.

4 Conclusion and Future Work

In this paper, a change detection method is proposed to detect the change in the image of earthen ruin. We use PCA to extract the relevant principal components and orthogonalize it to the vector space of each pixel of the difference image. Then we use the FCM algorithm to estimate the relationship between pixel and cluster center. The proposed method can effectively detect the change regions. We compare our method with other algorithms with different noise. The experiment results show that our proposed algorithm performs better than that of the compared methods.

Our future work will focus on the relation between change and disease factors to identify associated patterns of different diseases and image changes. Secondly, it is also a research direction to borrow deep learning methods to do trend prediction.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China (No. 61602379), international cooperation and exchange program of Shaanxi Province (No. 2016KW-034) and General industrial project of Shaanxi Province (No. 2018SF-369).

References

1. Mahdi, H., Teymoor, S.S.: Hyperspectral change detection: an experimental comparative study. Int. J. Remote Sens. **39**, 7029–7083 (2018)
2. Seydi, S.T., Hasanlou, M.: Sensitivity analysis of pansharpening in hyperspectral change detection. Appl. Geomatics **10**(1), 65–75 (2018)
3. Seydi, S.T., Hasanlou, M.: Sensitivity analysis of pansharpening in hyperspectral change detection. Appl. Geomatics **10**, 65–75 (2018)
4. Hussain, M., Chen, D., Cheng, A., Wei, H., Stanley, D.: Change detection from remotely sensed images: from pixel-based to object-based approaches. ISPRS J. Photogram. Remote Sens. **80**(2), 91–106 (2013)
5. Gang, C., Geoffrey, J.H., Luis, M.T.C., Michael, A.W.: Object-based change detection. Int. J. Remote Sens. **33**(14), 4434–4457 (2012)
6. Ghofrani, Z., Mokhtarzade, M., Reza Sahebi, M., Beykikhoshk, A.: Evaluating coverage changes in national parks using a hybrid change detection algorithm and remote sensing. J. Appl. Remote Sens. **8**(1), 1–16 (2014)
7. Celik, T.: Unsupervised change detection in satellite images using principal component analysis and k-means clustering. IEEE Geosci. Remote Sens. Lett. **6**(4), 772–776 (2009)
8. Lu, J., Li, J., Gang, C., Zhao, L., Xiong, B., Kuang, G.: Improving pixel-based change detection accuracy using an object-based approach in multitemporal SAR flood images. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **8**(7), 3486–3496 (2017)
9. Xiong, B., Chen, J.M., Kuang, G.: A change detection measure based on a likelihood ratio and the statistical properties of SAR intensity images. Remote Sens. Lett. **3**(3), 267–275 (2012)
10. Stephane, M., Charlotte, P.: Primal sketch of image series with edge preserving filtering application to change detection. In: Analysis of Multitemporal Remote Sensing Images. IEEE (2015)
11. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. **12**(7), 629–639 (1990)

12. Wang, G., Zhang, F., Jiao, L.: Change detection in SAR image based on multiscale product and PCA. In: Asian-Pacific Conference on Synthetic Aperture Radar. IEEE (2009)
13. Yan, D., Tang, L., He, L.: The microscopic image registration method based on parallel logarithmic subtraction template matching. In: Symposium on Photonics and Optoelectronics (2009)
14. Gao, T., Li, A., Meng, F.: Research on data stream clustering based on FCM algorithm 1. Procedia Comput. Sci. **22**, 595–602 (2017)
15. Bruzzone, L., Prieto, D.F.: Automatic analysis of the difference image for unsupervised change detection. IEEE Trans. Geosci. Remote Sens. **38**(3), 1171–1182 (2000)

Optimal Design Parameters for Tiny-YOLO2 Implementation with Light-Weight Embedded GPGPU Environment

Yechan Yu(✉), Daewoo Kim, Jinjoo Ha, Hojin Kim, and Kang Yi

School of Computer Science and Electrical Engineering,
Handong Global University, Pohang, Gyungbuk, Republic of Korea
{21500429, 21400067, 21000770, 21200213, yk}@handong.edu

Abstract. The aim of this paper is to find the optimal design for tiny YOLO2 in an embedded system environment. Our focus is to rebuild the given YOLO2 code and to find the optimal design parameters in order to maximize the speed using the light-weight GPGPU in a target embedded environment. To maximize the YOLO2 performance we used OpenCL framework while utilizing the embedded GPGPU and tried various aspects of OpenCL design parameters such work item, work group, and resulting in reducing the global memory access overhead and maximizing computing load balancing between computing units under constraints including local memory resources and computing resources. Our experimental results show that the overall performance enhancement is 18.2 times compared to the naive implementation.

Keywords: High-Performance Computing · OpenCL · YOLO · Embedded System · Parallel/Distributed System

1 Introduction

DCNN (deep convolutional neural network) is one of deep learning networks that is known to be suitable for object recognition based on image data [1, 2]. However, DCNN requires high computing resources and demands longer time and energy consumption even in prediction phase as well as training phase. Many researches have tried to solve this problem by using GPGPU and FPGA accelerators [3, 4]. One of recent trends in object recognition is that DCNN prediction is required in various embedded systems such as vehicles, mobile phone, and surveillance cameras to achieve edge computing where recognition should be performed on the nodes, not on the remote servers. In this paper, under the tight constraints from an embedded system, we have optimized the tiny YOLO2 [5] that is a popular DCNN open source for object detection and location. We tried to increase the memory efficiency with the limitation of a small amount of local memory in the embedded GPGPU system by utilizing blocking techniques [6, 7]. We implemented the methods to adjust the design parameters including workload for each computing units of embedded GPGPU to do fine tuning global memory and local memory accesses.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 36–43, 2020.
https://doi.org/10.1007/978-981-32-9244-4_5

2 Background

2.1 OpenCL

An OpenCL application consists of a host program and kernels. The host is a program that manages logic and data for parallel processing, and kernels are functions that execute parallel computations on device such as GPGPU. The host launches the kernels once it is ready to process data, and the kernels execute parallel computations. A PE (processing element) is the smallest physical unit that performs computation. Each work item is executed on the same PE, and every work item has its own private memory that other work items are not allowed to access. A CU (computing unit) consists of several PEs, and each work group is executed on the same CU. A work group has a local memory which all work items belonging to the same work group share. Global memory is an off-chip memory of device that allows access to all work items. This OpenCL memory model is illustrated in Fig. 1. One of GPU computing optimization techniques is to minimize the global memory accesses [8].

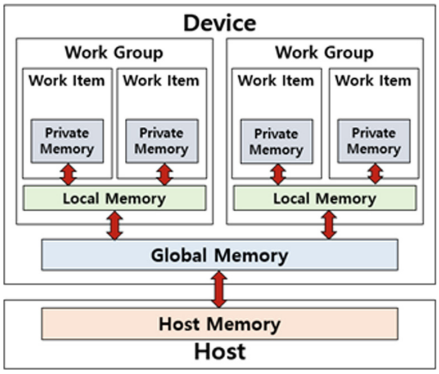


Fig. 1. OpenCL memory model

2.2 Our Target DCNN: YOLO Network

DCNN is a generally used deep learning model that is applied to object recognition. DCNN consists of many neural network layers, and each layer in DCNN detects features of objects. Due to the large number of layers and high complexity of each layer DCNN requires a number of parameters and needs a number of data movements between local and global memories. The application network used in this paper is the tiny YOLO [6] that has total of fifteen layers and has three types of layers: convolution, max-pooling, and fully connected. Tiny YOLO in this paper is implemented by "darknet" framework and can classify twenty different object types.

2.3 Our Target Board: R-Car H3

The embedded platform with GPGPU we used is R-car H3 from Renesas which targets automotive information processing. This platform has 10 CPU cores, IMG PowerVR GX6650 GPU supporting OpenCL v1.2. we implemented the tiny YOLO2 with OpenCL v1.2 on this embedded system. The program should be optimized in different way from PC environment because of the small amounts of local memory and CUs provided from this platform. Specifications of the board are shown in Table 1.

Table 1. R-car H3 properties.

Category	Properties
CPU	Quad ARM Cortex-A53, Quad ARM Cortex-A53 and ARM Cortex-R7 Dual lock-step
Memory	LPDDR4-3200 (Bandwidth 128b)
GPGPU	IMG PowerVR Series6XT GX6650
Number of GPGPU Core	192 Cores
Number of Local Memory	3068 Byte
Number of Compute Unit	6
Max Work Group Size	512

3 Accelerator Design Optimization Techniques

3.1 Blocking in Image Convolution

Efficient usage of memory and distribution of kernel's work are required to maximize the performance of GPU. In order to achieve the maximum performance of GPU, we used blocking method. Blocking is a method of finding the optimal memory structure size for memory reuse. Matrix multiplication algorithm using blocking method was introduced in [6, 7], and we used the blocking method for image convolution operations. Reusability of data is necessary due to the memory hierarchy, and blocking method allows re-usage of data in local memory.

Memory hierarchy of our computing platform consists of three levels: global memory, local memory, private memory. Figure 2 shows blocking method in a part of convolutional layer, and there are thirteen different parameters: X_i, Y_i, C, F_w, F_h, X_o, Y_o, K, WG_w, WG_h, WI_w, WI_h, w, which represents input width, input height, input channel number, Filter width, Filter height, output width, output width, output channel number, work-group width, work-group height, work-item width, work-item height, workload, respectively. Parameters X_i, Y_i, C have subscript 0, 1, and 2 which represent memory hierarchy. Parameters that are numbered with 0, 1, 2 uses private memory, local memory, and global memory, respectively. Equation (1) is an expression for

obtaining one output value of Convolution. The Kernel code was designed based on the equations in Eq. (1) to distribute the work to work-items. In addition, we set the work-group size according to the size of the local and private memory of the device and the size of output plane. We set the WI_w, WI_h and w variables to control the amount of the work given to a work-item using the equation in Table 2. The parameters used in Table 3 have a different value depending on the size of the Neural Network and the board used, and the larger the memory size, the higher the memory efficiency. Also, Fig. 3 shows the Hierarchical blocking of a single convolutional layer with the values in Table 2.

$$Output(k, y, x) = \sum_{c=0}^{C_2-1} \sum_{b=0}^{F_h-1} \sum_{a=0}^{F_w-1} input(k, c, y+b, x+a) \cdot weights(k, c, b, a), \qquad (1)$$

where $(0 \le x < X_o), (0 \le y < Y_o), (0 \le k < K)$.

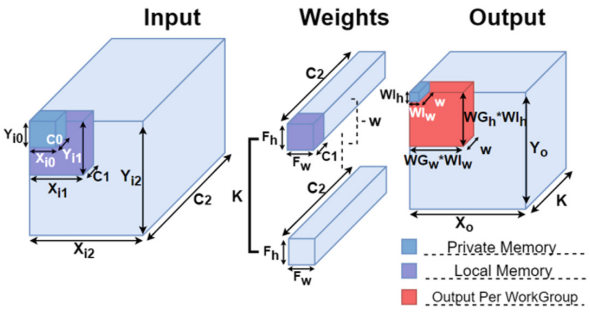


Fig. 2. Hierarchical blocking of a single convolutional layer

3.2 Optimal Workload

In the blocking method above, one work-item is responsible for the computation of one output node. However, we still have room for further optimization by making the global memory access more efficient. In the output image, computing the values at the same point on different planes requires exactly the same part of the input layer. In other words, the output node at (k_i, y, x) and (k_j, y, x) refer to the same address of the input layer for convolution operations. To take advantage of that, we utilized the concept of workload to avoid the redundant accesses to the input layer. In this method, each work-item, with given the workload w, is in charge of the computation of w output nodes. Figure 4 shows how this method differs from the previous one. With the workload concept, a work-item yields more than one output nodes as depicted in <case 2> of Fig. 4, while without the workload just one value of the output pixel comes out from a work-item as in <case 1> of Fig. 4. As a result, the application of the workload reduces not only the number of accesses to the input layer but also the number of work groups

Table 2. Equation of the parameter in Tiny-yolo v2.

Parameter	Size
X_{i0}	$(WI_w \cdot Stride + F_w - 1)$
X_{i1}	$(WG_w \cdot WI_w \cdot Stride + F_w - 1)$
Y_{i0}	$(WI_h \cdot Stride + F_h - 1)$
Y_{i1}	$(WG_h \cdot WI_h \cdot Stride + F_h - 1)$
Local memory	$(X_{i1} \cdot Y_{i1} + F_w \cdot F_h \cdot C1 \cdot w)$
Private memory	$(X_{i0} \cdot Y_{i0} + WI_w \cdot WI_h \cdot w)$

Table 3. Example of the parameter in Tiny-yolo v2

X_{i0}, Y_{i0}	X_{i1},Y_{i1}	C0, C1	Fw, Fh	WI_w,WI_h	WG_w, WG_h	w	X_{o1}, Y_{o1}
3	15	1	3	1	13	1	13

to 1/w of the original number of accesses and work groups, which leads to the alleviation of the memory traffic. However, as this method brings lower concurrency, it is necessary to find the optimal number of workloads depending on the specific device and the specific application.

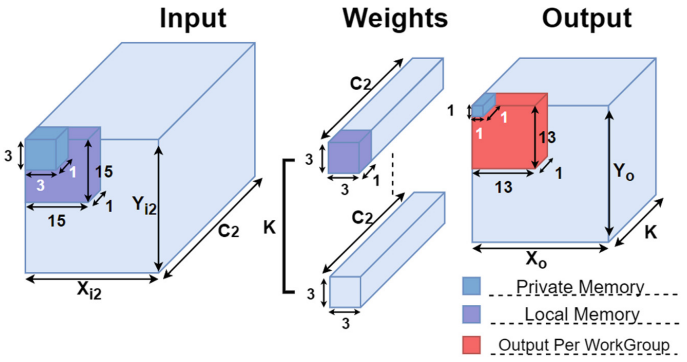


Fig. 3. Hierarchical blocking of a single convolutional layer in paper

4 Experimental Results

4.1 Performance Analysis Blocking Technique

We firstly applied the GEMM technique which is generally used for the acceleration of the convolution operations. However, it is only efficient in PC environment that has enough amount of memory. Since the amount of memory is small in embedded system, the image convolution with blocking method is faster because of less global memory access. Figure 5 shows the time ratio of image convolutions by blocking method with single core CPU, GEMM method with GPU acceleration, blocking method with GPU,

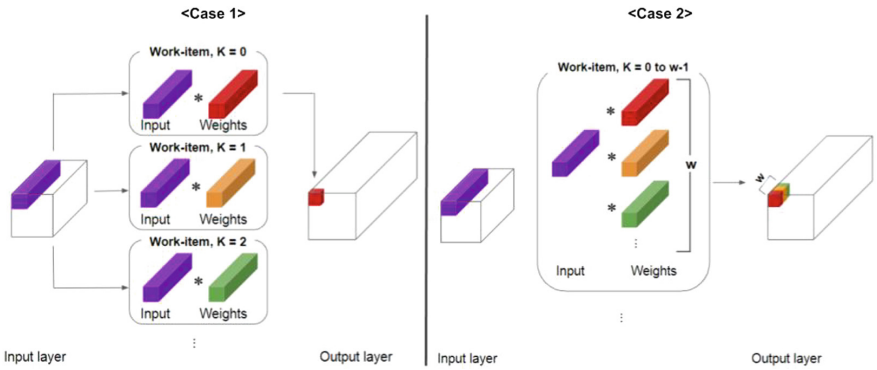


Fig. 4. Optimization using workload

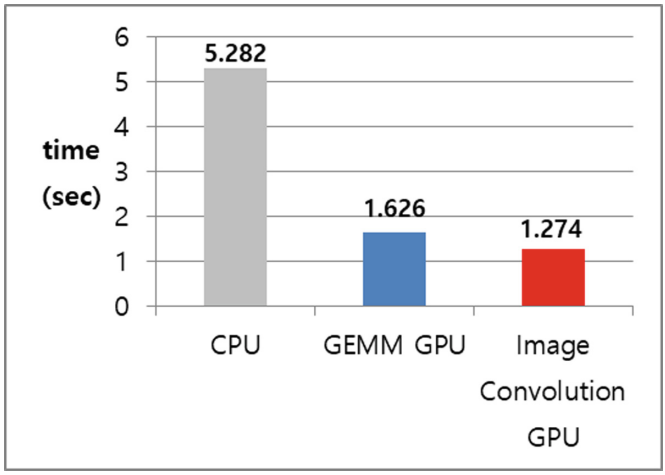


Fig. 5. Compare CPU, GEMM GPU and image convolution GPU time

respectively. Our image convolution operation by blocking with GPU is 3.25 times faster than CPU and 1.28 times faster than GEMM approach.

4.2 Exploring Design Space Further: Optimal Workload

The number of work groups, global memory accesses, and operations per work group should be considered in order to decide the workload per work item. In our experiment, all layers showed the maximum performance when the workload was eight. It is noticeable that the overall computing time graph is convex function curve with minima. We analyzed the reason of optimal point existence.

The change in the number of work groups and time consumed according to the change of workload in the 13-th layer of tiny YOLO2 is shown in Fig. 6. As the workload parameter increases, the number of work groups used in the kernel decreases,

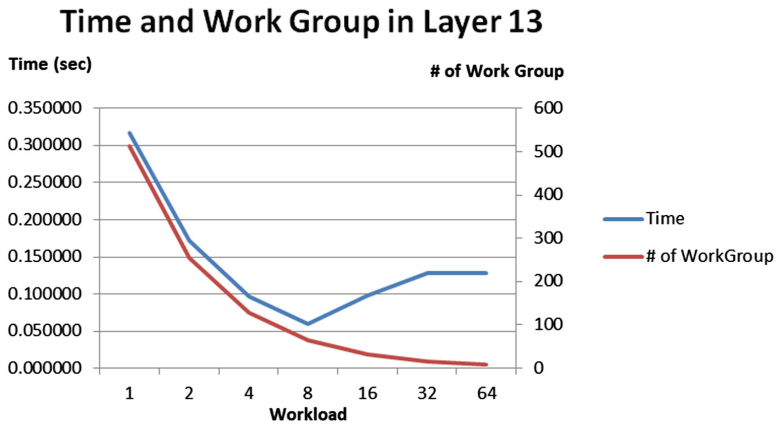


Fig. 6. Result time and number of work group in workload

and the overall kernel time decreases up to a certain point (workload = 8) and bounces back and slightly increase beyond the point.

The number of global memory accesses and the number of operations per work group are shown in Fig. 7. The number of global memory accesses required by the entire kernel is reduced by the inverse proportion of the workload, and the number of operations required per work group is increased accordingly. As the workload per work item increases, the number of work group decreases and the total memory access latency decreases. However, the amount of computations per work group increases. As the number of work groups decreases and the size of work group (number of work items per workgroup) increases, the CU may not be used efficiently due to the optimal computing load balance failure between CU. Therefore, the optimal number for the workload that balances the number of work groups and global memory access delay

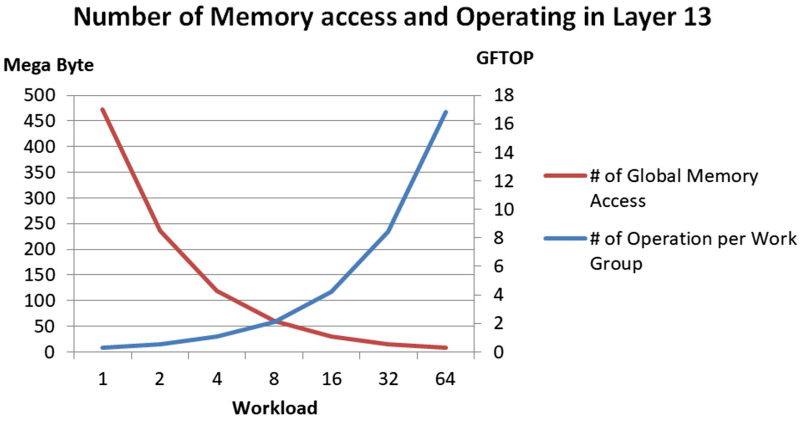


Fig. 7. Result memory access and number of operation in layer 13

time should be exist for each case. The embedded board we used had 6 CUs and found that the best performance is obtained when the workload is 8.

5 Conclusion

We applied blocking and workload change techniques to optimize DCNN under specific embedded GPGPU. Blocking is a method that uses small local and private memory efficiently. Image convolution using blocking method is 1.28 times faster than GEMM method. We also applied workload technique to reduce the number of global memory accesses and optimize the memory delay time. The number of work groups according to the number of CUs is also optimized. Using these technique, the overall computing speed was 18.2 times faster than the computing speed with single CPU.

Acknowledgments. This work was supported by National program for Excellence in software at Handong Global University (2017-0-00130) funded by Ministry of Science and ICT in Korea.

References

1. Farfade, S.S., Saberian, M.J., Li, L.J.: Multi-view face detection using deep convolutional neural networks. In: Proceedings of the 5th ACM on International Conference on Multimedia Retrieval, China, pp. 643–650 (2015)
2. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
3. Park, H., Lee, C., Lee, H., Yoo, Y., Park, Y., Kim, I., Kang, Y.: Optimizing DCNN FPGA accelerator design for handwritten hangul character recognition: work-in-progress. In: The 2017 International Conference (2017)
4. Elango, S.G.: Convolutional neural network acceleration on GPU by exploiting data reuse. San Jose State University (2017)
5. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016)
6. Matsumoto, K., Nakasato, N., Sedukhin, S.G.: Performance tuning of matrix multiplication in OpenCL on different GPUs and CPUs. In: 2012 SC Companion: High Performance Computing, Networking Storage and Analysis (2012)
7. Yang, X., Pu, J., Rister, B.B., Bhagdikar, N., Richardson, S., Kvatinsky, S., Kelley, J.R., Pedram, A., Horowitz, M.: A systematic approach to blocking convolutional neural networks. Stanford University (2016)
8. Tompson, J., Schlachter, K.: An Introduction to the OpenCL Programming Model. Person Education (2012)

Experiment Design and Analysis of Cross-Cultural Variation in Color Preferences Using Eye-Tracking

Bo Wu[1(✉)], Shoji Nishimura[1], Yishui Zhu[2], and Qun Jin[1]

[1] Faculty of Human Sciences, Waseda University, Tokorozawa, Japan
wubo@aoni.waseda.jp, {kickaha,jin}@waseda.jp
[2] School of Information Engineering, Chang'an University, Xi'an, China
yszhu@chd.edu.cn

Abstract. With the development of eye-tracking technology, existing studies verified the effect of culture on eye movements. However, the detail of how the different cultures, especially the different Asian cultures affect people's color preferences about products by visual attention remains in black box. In this paper, we focus on the difference between Chinese and Japanese who have similar but different cultures and try to identify their difference in visual attention about color preferences when selecting goods by using eye-tracking. The finding proves that people with different Asians cultures have different viewing patterns when faced to select goods with different colors. The results also indicate that culture as a factor does affect to ones' evaluation of products. The results can provide guidelines for products design or visual marketing management.

Keywords: Eye-tracking · Eye-movement analysis · Color preferences · Cultural variation

1 Introduction

Color is a powerful tool to attract a subject's attention, and to bring out the desire to consume [1]. Since people's color preference may affected by culture and geographical factors [2], many studies on color preference of culture have been done. However, prior studies usually see Asians to be holistic [3] and did not considered the difference of eye-movements among the people who similarly living in Asia, such as the Japanese and Chinese. Thus, in order to render the visual marketing efforts, it is important to identify whether the difference exists between these "similar foreigners" and locals in eye movements on product color design.

In this paper, we focus to identify the difference of Chinese and Japanese people in visual attention when selecting goods which used different color strategy by using eye-tracking. The scientific significance of our study is to provide a clear influence paths of different cultures on the color preferences by identifying the role of eye movements in this process. Different to the prior studies which focus on the difference between Westerners and Asians, this paper provides a smaller range to discuss the difference

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 44–49, 2020.
https://doi.org/10.1007/978-981-32-9244-4_6

between Chinese and Japanese who have similar but different cultures in a visual attention level. The use of eye-tracking data is helpful for managers to visualize the process of users' products selection, and the results of this study may provide guidelines for products design or visual marketing management.

2 Related Work

In recent years, based on the data collection by eye-tracking technology which enabled us to accumulate the detailed information on individuals' eye-based behaviors, existing studies verified the effect of culture on eye movements. For instance, based on the analysis of eye-movements data of American and Chinese participants while they viewed photographs with a focal object on a complex background, Chua et al. examined the possibility that the cultural differences arise from culturally different viewing patterns when they are confronted with a naturalistic scene [3]. However, this research sees Asians to be holistic and did not considered the difference of people who have similarly living in Asia, such as the Japanese and Chinese people who have similar but different cultures.

Color's effects on human performance and cognitive interpretation provide important evidence suggesting potential consumer reactions [4]. Prior studies find that color influences on consumers' related evaluation [5]. Many researchers had interest in the studies related to the different visual attention about colors of Asia people. For instance, based on the correspondence analysis of the choices of people from different area of Asia, Saito's research finds that different area (county) in Asia shows tendencies for unique color preference by using questionnaires [6]. The reasons given for the choices suggested that associative images based on environmental and cultural aspects may have important influences on color preference. On the other hand, according to the research of Lee et al. who utilized an eye-tracking experimental method to explore the possible relationships between color preferences and characteristics of scan-path [7]. Their result find that people's eye movement information does indicate to their color preference. However, they did not discussion about the culture effect and the specific effects on color preferences.

As a summary, prior researches indicated that color can influence people's evaluation about products, and according to the questionnaire survey, there are significant differences among the various Asian people who have different cultures with respect to the issues of color preferences. However, although previous studies have shown that the difference exists in eye movement between westerners and east Asians, the detail of how these different cultures, especially the different east Asian cultures affect customers' color preferences about products by eye-movements remains in black box. Therefore, based on the prior research and the actual situation of our experiment environment, the following two hypotheses are assumed.

H1. Subjects with different cultures have different eye movements when selecting goods with different color strategies.
H2. Subjects' culture as an influencing factor has effect people's product evaluation when selecting goods with different color strategies.

3 Experiment Design

3.1 Design Concept

We focus on the difference between Chinese and Japanese in a visual attention level when selecting goods. Based on the prior researches of eye-tracking in marketing field [8], a universal model for eye-movement related experiments is applied to this study. As shown in Fig. 1, according to this universal model, the influencing factors for people's eye-movement can be grouped into two types of factors, the top-down factors and the bottom-up factors. The top-down factors are these from subjects which will not change with the experiment environments, such as age and gender. On the other hand, the bottom-up factors are part of experiment environments, such as various color stimuli from products' pictures. Both types of factors will affect eye-movement metrics (average fixation duration, visit count, etc.), and finally lead to the downstream effects, which will be the score of color preference in this study.

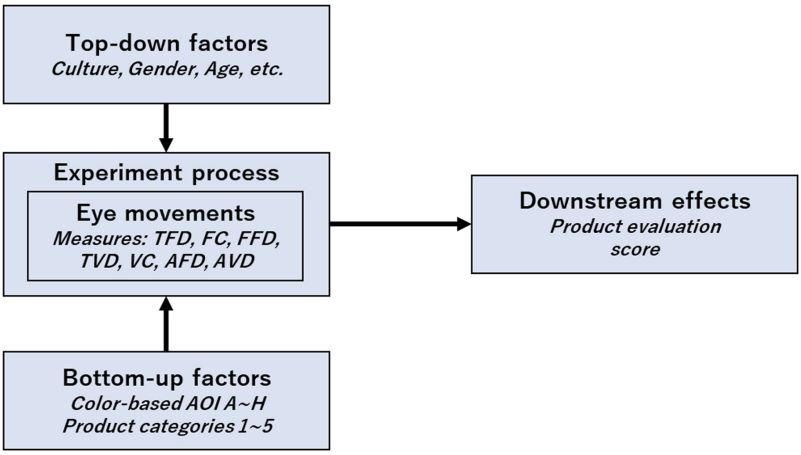


Fig. 1. The basic model of our experiment.

In this study, we plan to collect the eye-movement data from Chinese and Japanese customers who have similar but different cultures at an emulated in-store environment through laboratory experiments by using the eye-tracker device (Tobii Pro glass II). Ten Japanese college students and 30 Chinese overseas students are invited as the experimental participators to test their eye movements when facing to select goods with different color strategies.

3.2 Definitions of Measures

Except the culture and other demographic data such as gender and age, we defined the measures used to our experiment as follows.

Color Definition. To distinguish the effects of different colors, we used for data collection in the research were five kinds of product categories, include confectionery, cosmetic, clothing, medicine and electrical products which are kinds of goods categories that Chinese people like to buy when they travel to Japan [9]. In order to simulate the scene where customers choose products, each product was represented eight times in eight different NCS colors, include R, Y50R, Y, G50Y, G, B50G, B and R50B. Eight blocks in each experiment, and each block contains1 image of the product with a color applied on it. Each block will be given a different AOI (Area of interest) for eye-movement analysis. Their Numbers will be defined from A to H with corresponding virtual variables.

Eye-Movement Measures. For each AOI, based on the prior studies and our analysis software of our eye-tracker device, we use six measures to describe one's eye-movement. Include Total Fixation Duration (TFD), Fixation Count (FC), First Fixation Duration (FFD), Total Visit Duration (TVD), Visit Count (VC), Average Fixation Duration (AFD) and Average Visit Duration (AVD), which can be obtained directly through analysis software of Tobii II.

Product Evaluation Score (PEScore). After each experiment, subjects need to give a subjective ranking of preferences for 8 colors. The results will be calculated as a score to describe the evaluation to this type of color. For example, in one experiment, for the AOI "A" with color B (Blue), if the subjects' favorite top 1 color is also blue, the score for AOI "A" will be 1, which means the higher the score, the less the subject prefer this color.

4 Analysis and Discussion

In this section, to investigate the detail of how east Asian cultures affect subjects' color preferences and possible influencing factors, after the data cleansing and normalization process, related statistical hypothesis tests and regression analysis are presented. For the data collected by eye-tracker glasses (Tobii II) in the experiment, we got totally 2296 sets of records for all subjects, after pre-processing. The average time of each experiment record is almost 29.2 s.

Hypothesis H1 claims that subjects from different country may have different eye-movement measures when selecting goods with different color strategies. To test this hypothesis, Mann-Whitney U Test is applied as the non-parametric test.

As shown in Table 1, results indicate that between Chinese and Japanese, the difference of eye-movement measures are significant (p-value < 0.01), expect the measure of FFD (First Fixation Duration, p-value $= 0.553 > 0.05$). For most of eye-movement measures, Chinese people have a higher level than Japanese people (z-value < 0). Test result provides evidence for H1 that people have different East Asians cultures have different viewing patterns when faced to select goods with different color strategies. H1 is partly supported.

Table 1. Comparison of eye-movement measures.

	Group	N	Mean rank	Sum of	U	Z	P
TFD	Chinese	1770	1191.33	432731.50	315361.500	−8.904	.000**
	Japanese	484	894.07	2108653.50			
FC	Chinese	1770	1173.20	464824.50	347454.500	−6.389	.000**
	Japanese	484	960.38	2076560.50			
FFD	Chinese	1770	1131.75	538179.00	420809.000	−.594	0.553 n.s.
	Japanese	484	1111.94	2003206.00			
TVD	Chinese	1770	1188.44	437847.00	320477.000	−8.501	.000**
	Japanese	484	904.64	2103538.00			
VC	Chinese	1770	1177.87	456559.50	339189.500	−7.055	.000**
	Japanese	484	943.30	2084825.50			
AFD	Chinese	1770	1173.10	460683.50	343313.500	−6.638	.000**
	Japanese	484	951.83	2071691.50			
AVD	Chinese	1770	1161.34	483635.00	366265.000	−4.859	.000**
	Japanese	484	999.25	2053243.00			

*p < 0.05, **p < 0.01, n.s.: not significant

To test H2 hypothesizes, in order to obtain the specific influencing factors affecting subjects' products evaluation, based on the data we collected including demographic data (culture, gender, age, etc.), products categories (1–5) and color data (products color A–H) and eye-movements data (TFD, FC, FFD, TVD, VC, AFD, AVD), the forward stepwise regression analysis be applied with the PEScore as the dependent variable.

Table 2. Results of forward stepwise regression analysis.

Model	Unstandardized coefficients		Standardized coefficients	t	Sig.
	B	Std. error	Beta		
(Constant)	3.018	.143		21.140	.000
VC	1.839	.165	.237	11.124	.000
F	1.119	.147	.163	7.592	.000
C	.762	.146	.110	5.215	.000
D	.631	.150	.089	4.217	.000
A	−.772	.150	−.110	−5.160	.000
B	−.754	.145	−.110	−5.201	.000
Culture	.227	.110	.041	2.065	.039

As shown in Table 2, the results showed that the PEScore was estimated by VC, color A, B, C, D, F(Sig. < 0.01) and culture (Sig. < 0.05), r and R^2 were 0.397 and 0.158, respectively. Results shown that the factor of culture as an independent variable does affect to the PEScore, even though the effect is small, products may have a higher score (which means a lower preference) when the subjects come from Japan. H2 is

supported. In addition, VC as the factor have the largest weight, which means that if a subject will look at a product more than once when he/she did not prefer its color.

5 Conclusion

This paper provided a smaller range to discuss the difference between Chinese and Japanese who have similar but different cultures in a visual attention level when faced to select goods with different color strategies. Through statistical hypothesis tests and stepwise regression analyses, our study proved that except the measure of FFC, significant difference exists between Chinese and Japanese people on eye-movement measures. Our finding also proved culture as a factor has influence people's product evaluation. Moreover, our work finds that if a customer looks at products more than once, it may mean he/she did not prefer its color.

The finding of the study may be helpful to visualize the process of user's goods selection, and the results may help mangers to better understand the effects of color of products on different culture users which can provide guidelines for their products design or visual marketing management.

As for our future work, we will further improve the experiment with more influencing factors and try to use the technology of machine learning to predict customers' evaluation of the product based on their eye-movement measures.

Acknowledgments. This work was supported by 2018 Waseda University Grants for Special Research Projects with No. 2018S-166.

References

1. Lee, S., Barnes, H.: Using color preferences in magazine advertising. J. Advert. Res. **12**, 25–30 (1990)
2. Sakamoto, K: Cultural influence to the color preference according to product category. In: International Conference on Kansei Engineering and Emotion Research, pp. 1427–1523, Linköping (2014)
3. Chua, H.F., Boland, J.E., Nisbett, R.E.: Cultural variation in eye movements during scene perception. Proc. Natl. Acad. Sci. U.S.A. **102**(35), 12629–12633 (2005)
4. Babin, B.J., Hardesty, D.M., Suter, T.A.: Color and shopping intentions: the intervening effect of price fairness and perceived affect. J. Bus. Res. **56**(7), 541–551 (2003)
5. Crowley, A.E.: The two-dimensional impact of color on shopping. Mark. Lett. **4**, 59–69 (1993)
6. Saito, M.: A comparative study of color preferences in Japan, China and Indonesia, with emphasis on the preference for white. Percept. Mot. Skills **83**(1), 115–128 (1996)
7. Lee, T., Tang, D.L.: Exploring color preference through eye tracking. In: AIC Colour 2005, 10th Congress of the International Colour Association (2005)
8. Wedel, M., Pieters, R.: A review of eye-tracking research in marketing. In: Malhotra, N.K. (ed.) Review of Marketing Research, pp. 123–147. Emerald Group Publishing Limited, Bingley (2008)
9. Survey on consumption trend of foreigners visiting Japan. http://www.mlit.go.jp/kankocho/siryou/toukei/syouhityousa.html. Accessed 1 Apr 2019

Study on the Smart Speaker Security Evaluations and Countermeasures

Jiseop Lee, Sooyoung Kang, and Seungjoo Kim(✉)

Center for Information Security Technologies (CIST), Korea University, Seoul, South Korea
gsleegs4@gmail.com, skim71@korea.ac.kr

Abstract. The smart speaker provides users with useful functions such as music playback and online search with simple operation. However, since smart speakers always wait for the user's voice, if they are exposed to security threats, serious problems can occur such as eavesdropping and privacy disclosure. Therefore, in order to provide improved security for of all smart speakers, it is necessary to identify potential security threats and systematically investigate vulnerabilities. In this paper, we perform security threat modeling for four products with high market share. STRIDE threat modeling was used to make a checklist for systematic vulnerability checks and the checklist was used to check vulnerabilities of commercial devices. Here, we propose a new method to improve the security of smart speaker through the analysis of the vulnerability check result and the vulnerability of each product.

Keywords: Smart speaker · Threat modeling · STRIDE

1 Introduction

Smart speaker is a voice recognition platform that if a user delivers a voice command, artificial intelligence recognizes the voice and performs specific functions such as music playback, alarm setting, and Internet search. The smart speaker provides various useful functions and therefore, market size is growing rapidly [1]. However, smart speakers listen to the user's voice while waiting for input such as "Alexa" and "Ok Google", so if they are exposed to security threats, serious problems can occur such as eavesdropping and privacy disclosure. For example, recently, privacy was delivered from a person to another due to the malfunction of the smart speaker [2].

In this paper, we selected several security threat models suitable for smart speakers to make checklist [3] for systematic vulnerability checks. After checking the vulnerability of the commercial devices using the checklist, we propose how to improve the security level of smart speakers by case studies. Devices with high domestic and foreign market share due to high security level are chosen as models.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 50–70, 2020.
https://doi.org/10.1007/978-981-32-9244-4_7

2 Related Works

Threat modeling is a method to identify the overall security threats of a product using an organized structure. It is used to identify and remove vulnerabilities that may occur during the designing of the Security Development Lifecycle (SDL) to enhance security level. Additionally, it is one of an mitigation strategies for security threats to reduce latent damages [4]. We selected four well-known threat models and compared them to select suitable models for smart speakers. The contents are the same as Table 1.

Table 1. Comparison of threat modeling.

	STRIDE [5]	Trike [6]	LINDDUN [7]	PASTA [8]
Analysis focus	Design	Requirements	Design	Requirements
Analysis viewpoint	Software Vulnerability	Risk management for Asset	Privacy	Risk management for Business
Identification of system element	Data Flow Diagrams	Data Flow Diagrams & Use Flows	Data Flow Diagrams	Data Flow Diagrams
Threat determination methodology	STRIDE	Actor, Asset, Action matrix	LINDDUN	Threat-attack scenarios
Complexity	Medium	High	High	High
Documentation and support	Very good	Bad	Good	Good
Last update	2018	2012	2014	2015

All the threat modeling methods presented above classify and mitigate the threats, but each method approach to threats from different perspectives. However, STRIDE is the best approach to smart speakers if you consider analysis focus, analysis viewpoint, and documentation support for the threat model. It is well documented for security threats, and identifies security threats at the design step, and categorizes security threats from a S/W vulnerability viewpoint.

Meanwhile, The existing researches on improving security of smart speakers have been performed limitingly based on the personal experiences.

Hyunji Chung et al. described potential threats of Amazon Echo in four categories. The four categories contained disclosure of the firmware update of the wireless device, disclosure of a voice file to the cloud server, the sniffing of a transmitting data between the smart speaker and the server, and the risk of controlling the smart device through arbitrary voice command transmission [9]. However, this paper describes simple and general threats that are predictable, and studies on the security threats were not deep enough.

At DEFCON 26, the Tencent Blade Team proposed that for MIIO and Amazon Echo, possible threats can be occurred by disclosure of authentication tokens and firmware acquisition. In this study, by attacking the certain part of the smart speaker, possibility of a remote control attack was shown [10]. However, since this paper did not consider comprehensive security enhancement for the smart speakers, only limited solutions for specific function of the device were proposed.

3 Deriving Security Evaluation Criteria for Smart Speakers

3.1 Deriving Data Flow Diagram

We selected STRIDE as model system to identify threats in terms of vulnerability. To figure out the structure of the smart speaker DFD (Data Flow Diagram) was used. Since DFD shows the data flow of Entity, Process, Data Store, and Data Flow by visualization, it is easy to identify attack point and method if described accurately. The DFD description of smart speaker is shown in Fig. 1.

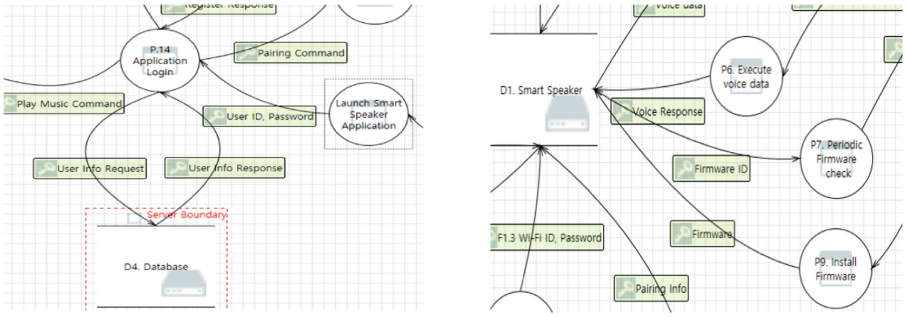


Fig. 1. Smart speaker data flow diagram

3.2 STRIDE Threat Modeling

STRIDE. It is possible to identify the threats of each component with STRIDE because the components of the DFD have their own potential threats. For a detailed explanation of STRIDE, refer to Appendix: STRIDE.

Attack Tree. To systemize all security threats, Attack Tree was visualized (Fig. 2). using Attack Library, a detailed threat lists for security threats and analysis targets, created with STRIDE. Refer to the Attack Tree, attack on smart speaker can be categorized into four types. Data acquisition, data modulation, denial of service attack, and speaker control.

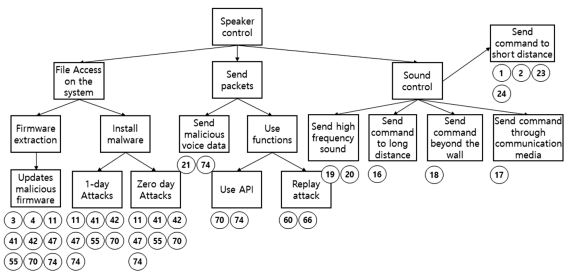


Fig. 2. The attack tree of the smart speaker

3.3 Vulnerability Checklist

Generally, when checking the vulnerability of a device, analysts or experts use their own know-hows or well-known checklist.

Table 2. A part of smart speaker checklist

Type	Surface	No	Detail	Threat No
Application	Authentication Policy	A1	Check number of login attempts limit	T11, T60, T61
		A2	Confirm password rule settings, including password length and special symbols, alphanumeric characters	T11
		A3	Confirm of privacy exposure	T6, T9, T58
		A4	Confirm of authentication related information (cookies, session values, tokens, etc.)	T5, T8, T10, T59
	APP Permission	A5	Confirm unnecessary permission requests for APP behavior	T70, T71
	Apply Encryption	A6	Check APK source code obfuscation	T70, T74
		A7	Communication data encryption check	T58, T59
Network	Port Scanning	N1	Open port check	T68, T71
		N2	Identify unnecessary management ports	T68, T71
	Packet Sniffing	N3	Confirm whether sensitive information is encrypted	T27, T32, T33, T70, T78, T86
		N4	Acquiring sensitive information through MITM	T6, T58, T69, T70
		N5	Confirm firmware acquisition	T33
	Packet Transmission	N6	Verifies speaker control through arbitrary commands or voice file transfer	T23, T24
		N7	Confirm replay attack possible (user account access, speaker function, etc.)	T23, T62
Hardware	Check Debugging port	H1	Check UART port	T68
		H2	Check JTAG port	T68
	Firmware Acquisition	H3	Firmware acquisition via UART/JTAG port	T38, T39, T71
		H4	Firmware acquisition through flash memory dump	T36, T39
	Firmware Modification	H5	Modify the firmware through the boot loader	T38
System	System shell Check	S1	Confirm ID, password exposures for administrators in the system shell	T68, T70
		S2	Confirmation of general user's access to administrator shell	T68, T70

However, since this approach cannot cover all areas of the object, it is difficult to improve security level. Therefore, it is necessary to make a new checklist specialized to the smart speaker in order to compensate for existing methods.

Attack vectors derived from STRIDE was distinguished into application, network, hardware and system. Additionally, in order to verify the completeness of the checklist, threats derived from each model was expressed in association with each checklist. The results are shown in Table 2.

Up to date, all threats of smart speakers have been identified according to systematic procedures such as DFD, STRIDE. Checklist of all smart speaker-related security threats and data collections being studied are included. These studies will provide researchers or security analysts a method of complete security check.

4 Security Evaluation of Smart Speakers

4.1 Vulnerability Assessment Using Security Evaluation Criteria

In this chapter, we used the checklist derived from Sect. 3 to evaluate the security level of A and B companies which are known to have high market share worldwide, and smart speakers of C and D companies with high market share in Korea. Then, weaknesses of each device have been compared and analyzed, and security threats are proposed.

The security evaluation environment was constructed similar to the actual one for the fairness and reliability of contents and methods. First, the smart speakers are updated to the latest firmware. The smart phone was installed the latest version of the application related to the smart speaker. Arbitrary environment such as installing malicious files and certificates were not considered.

Table 3. Evaluation results of application type

Type	No	A	B	C	D
Application	A1	O	O	O	O
	A2	O	O	O	O
	A3	X	O	X	O
	A4	O	O	O	O
	A5	O	O	O	O
	A6	O	O	O	O
	A7	O	O	O	O

Table 4. Evaluation results of network type

Type	No	A	B	C	D
Network	N1	O	O	O	O
	N2	O	O	O	O
	N3	O	O	O	O
	N4	O	O	O	X
	N5	O	O	O	X
	N6	O	O	O	O
	N7	O	O	O	O

Application. The evaluation results of the applications are shown in Table 3. The A1 and A2 were checked using the functions in the application, A3, A4, and A7 used data sniffing and Man In The Middle (MITM) attack [11], A5 and A6 were checked by static analysis of the application code.

(a) A3: During the log in process, ID and PW of user are exposed in the application of A, C. Figure 3 shows the result. In case of company A, all the contact information (name, phone number, etc.) are exposed when using Calling & Messaging. Figure 4 shows user contact information taken.

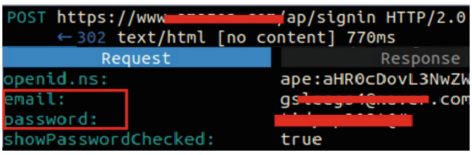


Fig. 3. Exposure of ID, Password

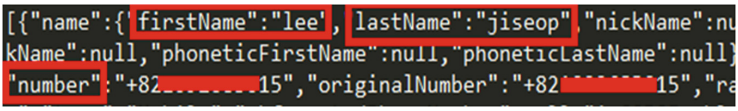


Fig. 4. User contact information

The application of each manufacturer was applied various protection techniques such as protection policy for privacy, code obfuscation, encryption communication using SSL, and pinning [12, 13]. However, some vendors have been vulnerable to MITM, which could threaten the disclosure of privacy such as user account information and contact information.

Network. The security evaluation results of networks are shown in Table 4. N1 and N2 used port scanning, N3 and N5 used sniffing, N4 used MITM, N6 used the APIs found from application code analysis, and N7 was checked by intercepting the packets and retransmitting them.

N4 is vulnerable to information disclosure due to MITM attack, and N5 is vulnerable to firmware update that is updated wirelessly. The results are as follows.

(a) N4: Device of company D can send a voice to smart speakers using application. By the vulnerability check, our group found that it was possible to obtain the voice data of the user transmitted from the application to the device D as shown in Fig. 5. If the user's voice is disclosed, privacy can be taken to invade privacy.
(b) N5: With sniffing, it was possible to get updating firmware of the company D's device (Fig. 6). Most of the smart speakers check the device's firmware version periodically to update the wirelessly update. However, if the firmware is disclosed, an attacker is able to eavesdrop and take out privacy.

By checking the network communication of the application-server and smart speaker-server, most devices from companies A to D have plans to counter security

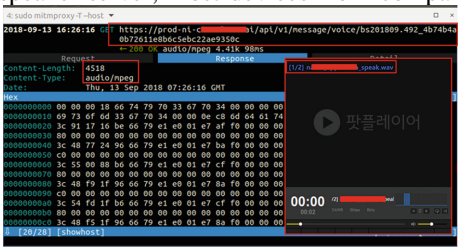


Fig. 5. Exposure of voice data

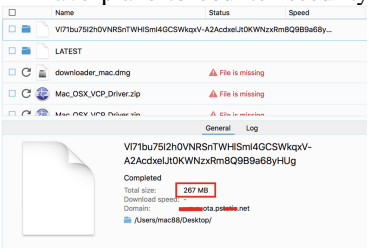


Fig. 6. Obtain firmware being wireless update

threats such as avoid using of unnecessary ports, encrypted communication and pinning using SSL. However, in case of company D, there was no encryption in certain functions. In addition, as security plan for MITM was absent, it is possible to eavesdrop and take out privacy.

Hardware. All the manufacturers removed means to access to UART and JTAG ports to prevent smart speaker debugging, firmware acquisition and modification. Therefore, checking every points of the system including the H4, H5 were impossible. As above, if debugging port is blocked, in all smart speakers, it becomes difficult to acquire and modify firmware, and access to the device file system will be difficult. Therefore, hardware-related attacks seem difficult.

4.2 Comparison and Analysis of Smart Speaker Vulnerabilities

As a result of security evaluation of smart speaker by four manufacturers, the potential vulnerabilities are shown in Table 5. In the checklist of applications, a vulnerability was found in the privacy disclosure (A3) of the authentication policy. In the networks, two vulnerabilities were found in packet sniffing MITM attack (N4) and firmware acquisition (N5). Most of the selected devices were vulnerable to sniffing and MITM attacks among the network attacks.

Table 5. Compare smart speaker vulnerabilities

Checklist	A	B	C	D
Application A3	Account and Contact information	X	Account information	X
Network N4	X	X	X	Voice file exposure
Network N5	X	X	X	Firmware acquisition

5 Conclusion

In conclusion, we have derived a new security evaluation criteria for smart speakers using STRIDE to evaluate the security level of commercial devices. In selected smart speakers, disclosure of privacy and disclosure of firmware and voice files of wireless updating were found. Based on the results our group compare and analyse potential threats of smart speakers and found MITM attack is possible in most devices.

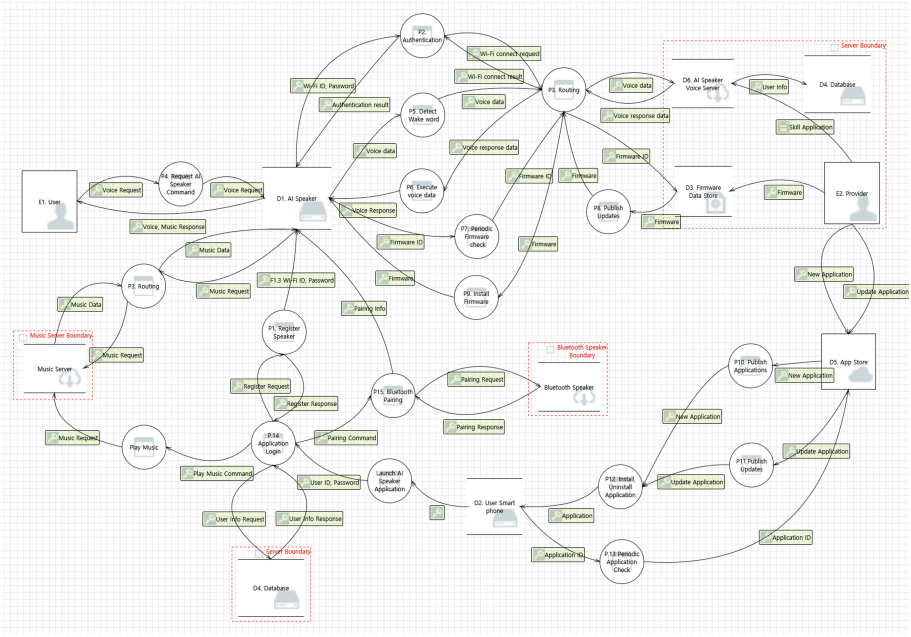
MITM attacks are tricking MAC and IP to intercept and listen to packets between users and servers which can lead serious problems such as disclosure of privacy, contacts or voices. Therefore, if smart speakers are applied SSL encryptions, certificate and public key pinning will minimize MITM attacks. Although smart speakers have different functions for each manufacturer, they contain common functions related to

user's privacy such as voice recognition, scheduling, alarming, and memo. By utilizing security evaluation standards proposed in this paper, it could be possible to make a checklist applicable to specific smart speakers.

Acknowledgements. This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2017-0-00184, Self-Learning Cyber Immune Technology Development).

Appendices

Data Flow Diagram



Attack Library

No	Year	Type	Source/Author	Title
1	2011	Conference	Black Hat 2011/Joe Grand	Hardware Reverse Engineering: Access, Analyze,& Defeat
2	2013	Conference	DEFCON 21/Phorkus and Evilrob	Hacking Embedded Devices, Doing Bad Things to Good Hardware
3	2013	Conference	Black Hat 2013/J Zaddach	Embedded Devices Security and Firmware Reverse Engineering
4	2014	Conference	Australian Information Security Management Conference/Veelasha Moonsamy, Lynn Batten	Mitigating man-in-the-middle attacks on smartphones – a discussion of SSL pinning and DNSSec
5	2015	Conference	Zeronights/Alexander, Boris	Practical Firmware Reversing and Exploit Development for AVR-based Embedded Devices
6	2015	Conference	International Telemetering Conference/Wondimu Zegeye, Richard A Dean, Farzad Moazzami, Yacob Astatke	Exploiting Bluetooth Low Energy Pairing Vulnerability in Telemedicine
7	2017	Conference	Black Hat 2017/Ben Seri and Alon Livne	Exploiting BlueBorne in Linux-based IoT devices
8	2017	Conference	HITBSecConf/Slawomir Jasek	Blue picking – hacking Bluetooth Smart Locks
9	2018	Conference	DEFCON 26/Tencent Blade Team	Breaking Smart Speaker, We are Listening to you
10	2017	Conference	Black Hat 2017/Sen Nie, Ling Liu, Yuefeng Du	Free-Fall: Hacking Tesla from Wireless to CAN BUS
11	2016	Conference	Black Hat 2016/Chilik Tamir	Su-a-Cyder: Home-Brewing iOS Malware Like a B0$$
12	2011	Conference	28th Chaos Communication Congress/Dario Carluccio	Smart Hacking for Privacy
13	2018	Vulnerability	CVE	CVE-2018-9070
14	2016	Web document	Cert Italia	Malware Android “CALLJAM” SCOPERTO SU GOOGLE PLAY
15	2016	Web document	Cert Italia	“DRESSCODE”: NUOVO MALWARE ANDROID SCOPERTO SU GOOGLE PLAY
16	2014	Web document	DistriNet	LINDDUN: Privacy Threat Modeling
17	2017	Paper	University of Notre Dame/Yuan Gong, Christian Poellabauer	Crafting Adversarial Examples for Computational Paralinguistic Applications

(*continued*)

(*continued*)

No	Year	Type	Source/Author	Title
18	2010	Paper	IEEE Design & Test of Computers	Attacks and Defenses for JTAG
19	2013	Paper	NDSS 2013/Ang Cui, Michael Costello, Salvatore J. Stolfo	When Firmware Modifications Attack: A Case Study of Embedded Exploitation
20	2014	Paper	2014 IEEE Symposium on Security and Privacy/Lin Shung Huang, Alex Rice, Erling Ellingsen	Analyzing Forged SSL Certificates in the Wild
21	2015	Paper	IJRSCSE 2015/Vimalesh Kumar Dubey, Kumari Vaishali, Nishant Behar, Manish Shrivastava	A Review on Bluetooth Security Vulnerabilities and a Proposed Prototype Model for Enhancing Security against MITM Attack
22	2016	Paper	USENIX 2016/Nicholas Carlini, Pratyush Mishra, Tavish Vaidya, Yuankai Zhang, David Wagner	Hidden Voice Commands
23	2017	Paper	IEEE Computer Society/Hyunji Chung, Michaela loarga, Jeffrey Voas, Sangjin Lee	Alexa, Can I Trust You?
24	2017	Paper	ACM Conference on Computer and Communications Security (CCS)/Xiaoyu Ji	DolphinAttack: Inaudible Voice Commands
25	2017	Paper	MIT/William Haach, Michael Wallace	Security Analysis of the Amazon Echo
26	2017	Paper	Michigan State University/Xinyu Lei, Guan-Hua Tu, Alex X. Liu	The Insecurity of Home Digital Voice Assistants - Amazon Alexa as a Case Study
27	2017	Paper	ACM CCS 2017/Mathy Vanhoef	Key Reinstallation Attacks: Forcing Nonce Reuse in WPA2
28	2017	Paper	Personal and Ubiquitous Computing/Da-Zhi Sun, Yi Mu, Willy Susilo	Man-in-the-middle attacks on Secure Simple Pairing in Bluetooth standard V5.0 and its countermeasure
29	2017	Paper	Australian Information Security Management Conference/Brian Cusack, Bryce Antony, Gerard Ward	Assessment of security vulnerabilities in wearable devices
30	2014	Paper	2014 ACM SIGSAC Conference on Computer and Communications Security/Yeongjin Jang, Chengyu Song, Simon P. Chung, Tielei Wang, Wenke Lee	A11y Attacks: Exploiting Accessibility in Operating Systems

(*continued*)

(*continued*)

No	Year	Type	Source/Author	Title
31	2018	Paper	Deep Learning and Security Workshop/Nicholas Carlini, David Wagner	Audio Adversarial Examples: Targeted Attacks on Speech-to-Text
32	2018	Paper	NIPS 2017 Machine Deception workshop/Moustafa Alzantot	Did you hear that? Adversarial Examples Against Automatic Speech Recognition
33	2018	Paper	Beijing Key Laboratory of IoT Information Security Technology/Nan Zhang, Xianghang Mi, Xuan Feng	Understanding and Mitigating the Security Risks of Voice-Controlled Third-Party Skills on Amazon Alexa and Google Home
34	2018	Paper	Rongjunchen Zhang, Xiao Chen, Jianchao Lu	Using AI to Hack IA: A New Stealthy Spyware Against Voice Assistance Functions in Smart Phones
35	2018	Paper	ACMSE 2018/Richmond, Kentucky	Testing vulnerabilities in Bluetooth Low Energy
36	2017	Paper	IEEE Access/Efthimios Alepis, Constantinos Patsakis	Monkey Says, Monkey Does: Security and Privacy on Voice Assistants
37	2015	Paper	IEEE Transactions on Electromagnetic Compatibility/Chaouki Kasmi, Jose Lopes Esteves	IEMI Threats for Information Security: Remote Command Injection on Modern Smartphones
38	2015	Paper	UseNIX 2015/Tavish Vaidya, Yuankai Zhang, Micah Sherr, Clay Shields	Cocaine Noodles: Exploiting the Gap between Human and Machine Speech Recognition
39	2017	Paper	CSAE 2017/Xiao Fu, Zhijian Wang, Yong Chen, Feng Ye	Research on Android Application Package Stealth Download Hijacking
40	2018	Paper	Ben Gurion University/Or Ami, Yuval Elovici, Danny Hendler	Ransomware Prevention using Application Authentication-Based File Access Control
41	2017	Paper	ICISS 2017/Anis Bkakria, Mariem Graa, Nora Cuppens-Boulahia	Experimenting Similarity-Based Hijacking Attacks Detection and Response in Android Systems
42	2012	Paper	SEC 2012/Alessandro Armando, Alessio Merlo, Mauro Migliardi, Luca Verderame	Would you mind Forking This Process? A Denial of Service attack on Android
43	2014	Paper	SPSM 2014/Steven Arzt, Stephan Huber, Siegfried Rasthofer, Eric Bodden	Denial-of-App Attack: Inhibiting the Installation of Android Apps on Stock Phones
44	2017	Paper	ASIA CCS 2017/Behnaz Hassanshahi, Roland H.C. Yap	Android Database Attacks Revisited

(*continued*)

(*continued*)

No	Year	Type	Source/Author	Title
45	2017	Paper	IJCSMC 2017/Zainab S. Alwan, Manal F. Younis	Detection and Prevention of SQL Injection Attack: A Survey
46	2017	Paper	2017 IEEE Symposium on Security and Privacy/Nethanel Gelernter, Senia Kalma, Bar Magnezi, Hen Porcilan	The Password Reset MitM Attack
47	2017	Paper	ISC 2017/XingXing Wang	Improving Password Guessing using Byte Pair Encoding
48	2015	Paper	IJCSIT 2015/Aqib Malik	A Model to Restrict Online Password Guessing Attacks
49	2018	Paper	IEEE 2018/Roberto Merco	Replay Attack Detection in a Platoon of Connected Vehicles with Cooperative Adaptive Cruise Control
50	2017	Paper	INTERSPEECH 2017/Parav Nagarsheth	Replay Attack Detecting using DNN for Channel Discrimination
51	2017	Paper	2017 IEEE International Symposium on Circuits and Systems/Mohammad Raashid Ansari	A low-cost masquerade and replay attack detection method for CAN in automobiles
52	2018	Paper	Georgetown University Law Center/David A. Hyman	Implementing Privacy Policy: Who should Do What?
53	2015	Paper	Journal of Computer and Security/Nader Sohrabi Safa	Information security policy compliance model in organizations
54	2017	Paper	23rd ACM SIGKDD/Lu Zhang	Achieving Non-Discrimination in Data Release
55	2015	Paper	Privacy Enhancing Technologies/Reza Shokri	Privacy Games: Optimal User-Centric Data Obfuscation
56	2017	Paper	Cryptography and Security/Samuel Yeom	Privacy Risk in Machine Learning: Analyzing the Connection to Overfitting
57	2017	Paper	International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics/Ashalatha R	Data storage security algorithms for multi cloud environment
58	2015	Paper	International Conference on Next Generation Computing Technologies/Preeti Sirohi	Cloud computing data storage security framework relating to data integrity, privacy and trust
59	2014	Paper	Intelligent Information Hiding and Multimedia Signal Processing/Jen Ho Yang	An ID-Based User Authentication Scheme for Cloud Computing

(*continued*)

(*continued*)

No	Year	Type	Source/Author	Title
60	2016	Paper	IEEE Signal Processing Magazine/Vishal M. Patel	Continuous User Authentication on Mobile Devices: Recent progress and remaining challenges
61	2012	Technical Report	SANS/Neil Jones	Exploiting Embedded Devices
62	2012	Technical Report	Inverse Path/Andrea Barisani, Daniele Bianco	Practical Exploitation of Embedded Systems
63	2015	Technical Report	Samsclass/Sam Bowne	Making an SSL Auditing Proxy with a Mac and Burp
64	2016	Technical Report	Vanderpot/Ike Clinton, Lance Cook, Dr. Shankar Banik	A Survey of Various Methods for Analyzing the Amazon Echo
65	2016	Technical Report	Oxford University	Security Vulnerabilities in Speech Recognition Systems
66	2016	Technical Report	Vanderpot	Amazon Echo Rooting: Part 1
67	2016	Technical Report	Vanderpot	Amazon Echo Rooting: Part 2
68	2017	Technical Report	MWR labs/Mark Barnes	Alexa, are you listening?
69	2017	Technical Report	medium.com/micaksica	Exploring the Amazon Echo Dot, Part 1: Intercepting firmware updates
70	2017	Technical Report	medium.com/micaksica	Exploring the Amazon Echo Dot, Part 2: Into MediaTek utility hell
71	2017	Technical Report	NowSecure/Rono Dasgupta	Certificate pinning for Android and iOS: Mobile man-in-the-middle attack prevention
72	2017	Technical Report	Securing/Slawomir Jasek	Gattacking Bluetooth Smart Devices
73	2018	Technical Report	CanSecWest/HyperChem	Practical JTAG: From 0 to 1
74	2015	Technical Report	Charlie Miller, Chris Valasek	Remote Exploitation of an Unaltered Passenger Vehicle
75	2018	Technical Report	IEEE Security & Privacy/Lee Garber	Security, Privacy, Policy, and Dependability Roundup

STRIDE

Type	No	Name		Threat description	Attack library	Threat No
Entity	E1	User	S	The attacker masquerades as a User	18, 24, 25, 27	T1
			R	The attacker denies the control of the speaker	18, 24, 25, 27	T2
Entity	E2	Provider	S	The attacker masquerades as a Provider	36, 45	T3
			R	The attacker denies the provided of the speaker	36, 45	T4
Process	P1	Register Speaker	T	Threats to manipulate authentication values in transit	4, 10, 13, 16, 31, 57	T5
			I	Threats that expose User's ID, Password	4, 13, 16, 22, 31, 43	T6
			D	Threats that prevent you from performing User authentication by passing invalid argument values	4, 10, 31, 43, 57	T7
Process	P2	Authentication	T	Threats to manipulate authentication values in transit	4, 13, 20, 31, 43, 57	T8
			I	Threats that expose User's ID, Password	4, 13, 20, 22, 31, 43	T9
			D	Threats that disable data transmission of authentication values	4, 20, 31, 43	T10
			E	Threats to gain the privileges of a router through a specific attack	20, 29, 30	T11
Process	P3	Routing	S	After an attacker obtains an administrator account using a random assignment attack, Attacker masquerades as a administrator	29, 30	T12
			T	Threats to tamper with existing file system files	29, 30, 45	T13
			R	Threats denying actions such as accessing, running, or tampering with the file system	29, 30, 43, 45	T14
			I	Threats that expose users' data flowing through the router	4, 13, 20, 22, 31, 43	T15

(*continued*)

(*continued*)

Type	No	Name		Threat description	Attack library	Threat No
Process	P4	Request Smart Speaker Command	S	Disguised as user through long distance voice command	16, 18	T16
			S	Disguised as a user using voice commands over a communication medium	16, 18	T17
			S	Beyond the wall, disguised as a user through voice commands	16, 18	T18
			S	Disguised as user through high frequency voice attack	15, 17, 18, 24, 25, 39, 46	T19
			R	Threat of denying high frequency voice attacks	15, 17, 18, 24, 25, 39, 46	T20
			I	Threats that expose your information through malicious voice	6, 18, 23, 24, 26, 39, 46	T21
			D	Threats that cannot use speakers through malicious voice	6, 18, 23, 24, 39, 46	T22
Process	P5	Detect Wake Word	S	An attacker masquerades as a user and instructs the speaker	6, 16, 18, 23, 24, 25, 26, 27, 39, 46	T23
			R	Threat of denying of attacker's voice	6, 16, 18, 23, 24, 25, 26, 27, 39, 46	T24
Process	P6	Execute Voice Data	S	Using IP spoofing to trick speakers into disguise as a server	4, 13, 23, 31	T25
			T	Threats to send manipulated voice files to speakers	4, 13, 23	T26
			R	Threats that an attacker denies camouflage and tampering	4, 13, 23	T27
			D	Threats that do not use speakers through manipulated voice files	6, 23, 24	T28
Process	P7	Periodic Firmware Check	T	Threats to modify firmware information	4, 5, 12, 38, 43, 49	T29
			I	Threats that expose firmware information	4, 5, 12, 38, 43, 49	T30
			D	Threats that make the update impossible by modifying version information	12, 43, 49	T31

(*continued*)

(*continued*)

Type	No	Name		Threat description	Attack library	Threat No
Process	P8	Publish Firmware Updates	S	Malicious firmware disguised as normal firmware	1, 5, 12, 38, 49	T32
			T	Threats to modify firmware to update	1, 3, 12, 38, 43, 49	T33
			R	Threat to denial of tampered firmware installation	1, 12, 38, 43, 49	T34
			I	Threats that expose the system information contained in the firmware	5, 12, 38, 41, 49	T35
			I	Threats to fetch firmware files	1, 5, 12, 38, 41, 43, 49	T36
			D	Threats to disable firmware updates	12, 43, 49	T37
Process	P9	Install Firmware	T	Threats to modify firmware	1, 3, 5, 10, 11, 32, 34, 35, 40, 41, 42, 45, 49	T38
			R	Threat to deny firmware tampering and installation	1, 3, 10, 32, 34, 35, 40, 41, 42, 45, 49	T39
			D	threat that disables system operation by installing corrupted firmware	1, 3, 5, 10, 11, 32, 34, 35, 40, 42, 45, 49	T40
			E	Threat that an attacker installs a malicious file	1, 3, 5, 10, 11, 32, 34, 35, 40, 41, 42, 45, 49	T41
			E	Threat that user's voice eavesdropping through modulated firmware	5, 12, 38, 49	T42
Process	P10	Publish Applications	S	Disguising a malicious application as a normal application	12, 36, 50, 51, 52, 54	T43
			T	Threats to tamper with installed applications	36, 50, 52, 54	T44
			R	Threat to denial of application installation history	36, 50, 54	T45
			I	Threats that expose information in the application	36, 50, 54	T46
			E	Threats that control smartphone functionality through malicious applications	36, 47, 48, 51	T47

(*continued*)

(*continued*)

Type	No	Name		Threat description	Attack library	Threat No
Process	P11	Publish Application Updates	S	Malicious application disguised as normal application	12, 36, 50, 51, 52, 54	T48
			T	Threats to tamper with updated applications	12, 50, 52, 54	T49
			R	Threat to denial of application tampering	12, 50, 52, 54	T50
			I	Threats that expose information in application	36, 50, 54	T51
			D	Threats that disable application updates	12, 50, 54	T52
Process	P12	Install, Uninstall Application	T	Threats to install moderated files	36, 38, 50, 51	T53
			D	Threats that disrupt system operation by installing corrupted files	36, 51, 54	T54
			E	Threats that unauthorized users install files	36, 51, 54	T55
Process	P13	Periodic Application Check	T	Threats to tamper with application information	4, 13, 43	T56
			I	Threats that expose application information	4, 13, 22, 31, 43, 54	T57
			D	Threats that make the update impossible by modifying version information	4, 13, 31, 43, 54	T58
Process	P14	Application Login	S	masquerades as a user through password guessing attack	59, 60	T59
			S	Disguised as user through replay attack	61, 62, 63	T60
			I	Threats that exposed authentication values are exposed	4, 13, 31, 53, 54	T61
			I	Threats that expose other information needed for additional authentication	4, 13, 22, 31, 53, 54	T62
			D	Threats that make login page inaccessible	4, 13, 31, 53	T63
			D	Threats that exceed the number of login attempts and make normal access impossible	4, 53	T64
			E	Threats accessible to user accounts using exposed authentication values	4, 13, 54	T65

(*continued*)

(*continued*)

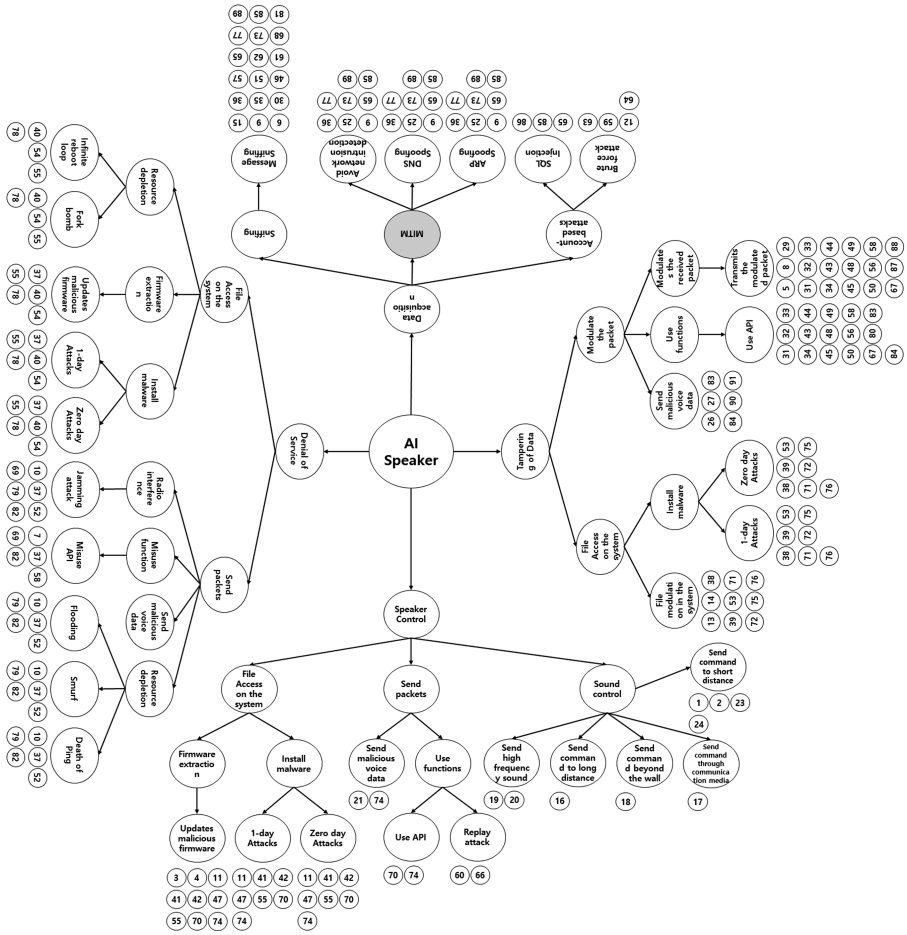
Type	No	Name		Threat description	Attack library	Threat No
Process	P15	Bluetooth Pairing	S	Disguised as user through replay attack	61, 62, 63	T66
			T	Threat of modifying transmission information of Bluetooth	7, 9, 14, 21	T67
			I	Threats that Bluetooth transmission information is exposed	7, 9, 14, 21, 28	T68
			D	Threat that Bluetooth pairing impossible	9, 14, 21	T69
			E	Threats to gain the rights of a device through a particular attack	8, 44	T70
Data Store	D1	Smart Speaker	T	Speaker system files, threats to tamper with memory	1, 3, 5, 10, 11, 32, 34, 35, 40, 41, 42, 45, 49	T71
			R	The threat that an attacker controls the device and denies this behavior	6, 16, 18, 19, 24, 25, 33, 39, 46	T72
			I	Threats that expose sensitive information (voice, schedule, etc.)	6, 16, 18, 19, 23, 24, 25, 33, 34, 35, 37, 39, 46	T73
			D	threat that prevents a device from being used for a certain amount of time	6, 10, 16, 23, 24, 25, 32, 34, 35, 39, 40, 42, 53	T74
Data Store	D2	User Smart phone	T	a threat that modifies memory on smartphone	1, 2, 29, 30	T75
			R	Threats denying access to system files memory on smart phones	1, 2, 29, 30	T76
			I	Threats that expose system files, memory contents on smart phones	1, 2, 3, 29, 30	T77
			D	Threats that do not provide service due to memory corruption	29, 30, 53, 54	T78
			D	Threats that fail to provide services due to network paralysis	29, 30, 53, 54	T79

(*continued*)

(*continued*)

Type	No	Name		Threat description	Attack library	Threat No
Data Store	D3	Firmware Data Store	T	Threats to manipulate transmitted firmware	12, 43, 49, 58	T80
			I	Threats that expose firmware data	10, 12, 16, 18, 41, 43, 49, 58	T81
			D	Threats not uploading firmware	12, 43, 58	T82
Data Store	D4	Database	T	Threats to transmit modulated data	4, 10, 31, 43	T83
			R	Threat denying modulated data transmission	4, 31, 43	T84
			I	Threats that expose your sensitive information	4, 10, 19, 22, 31, 43	T85
			D	Threats that fail to provide services through arbitrary queries	10, 55, 56	T86
Data Store	D5	App Store	T	Threat to upload moderated apps to the App Store server	47, 48	T87
			R	Threats denying access to the App Store server	47, 48	T88
			I	Threats to exposure of application data	19, 22	T89
Data Store	D6	Smart Speaker Voice Server	T	The threat that an attacker sends a malformed voice file to the server	18, 24, 25	T90
			R	A threat that denies attackers from sending malformed voice files to the server	18, 24, 25	T91

Attack Tree



References

1. S&P Global Market Intelligence. https://www.spglobal.com/marketintelligence/en/news-insights/research/smart-speakers-take-off
2. CNBC. https://www.cnbc.com/2018/05/24/amazon-echo-recorded-conversation–sent-to-random-person-report.html
3. Chun Yu, C.: Threat modeling techniques. Technical report, Delft University of Technology (2016)
4. Adam, S.: Threat Modeling. Wiley, Hoboken (2014)
5. Microsoft Azure. https://docs.microsoft.com/ko-kr/azure/security

6. Trike. www.octotrike.org
7. LINDDUN Privacy threat modeling. https://linddun.org/
8. Tony, U., Marco, M.: Risk Centric Threat Modeling: Process for Attack Simulation and Threat Analysis. Wiley, Hoboken (2015)
9. Hyunji, C., Michaela, L., Jeffrey, V., Sangjin, L.: Alexa, Can I Trust You? In: IEEE Computer, pp. 100–104. IEEE Computer Society (2017)
10. Tencent Blade Team: Breaking Smart Speaker- We are Listening to you, DEFCON 26, Las Vegas (2018)
11. Blackhat. https://www.blackhat.com/presentations/bh-europe-03/bh-europe03-valleri.pdf
12. OWASP. https://www.owasp.org/index.php/Certificate_and_Public_Key_Pinning
13. Mahesh, B., Deepak, K.: Certificate pinning for Android applications. In: 2017 International Conference on Inventive Systems and Control, IEEE, Coimbatore (2017)

Research on Tilt Survey Algorithm Based on Accelerometer and GNSS Receiver Data Fusion

Caixu Xu[1] and Jie He[2(✉)]

[1] College of Software, Central South University, Changsha, China
xucaixu0815@163.com
[2] Guangxi Colleges and Universities Key Laboratory of Image Processing and Intelligent Information System, Wuzhou University, Wuzhou, China
hejiel213@126.com

Abstract. At present, in the process of providing accurate location service for high-precision GNSS receiver of tilt survey, external environment, hardware structure and other factors will restrict the accuracy, universality and stability of service. To solve these problems, a method is proposed an improved attitude tilt survey algorithm only based on data fusion solution of an accelerometer and a GNSS receiver. On the premise of not changing hardware of the GNSS receiver, the algorithm enhances stability and universality of the attitude tilt survey. In the end, the measurement environment was chosen for comparative verification: Results indicate that the algorithm is slightly higher than attitude tilt survey proposal based on magnetic compass system in position correction accuracy, eliminates complicated calibration and preparation of sensors, reduces cost of operations and materials, As verified, the algorithm has outstanding advantages and wide application outlook in the field of precision engineering position services.

Keywords: Position service · Data fusion · Tilt survey · Integrated navigation

1 Introduction

With the development of GNSS (Global Navigation Satellite System) and MEMS (Micro Electro Mechanical System) sensors, real-time, intelligent and high-accuracy positioning [1] for LBS (Location Based Services) is widely applied to fields such as engineering staking-out, deformation monitoring, ground feature prospecting, etc. the GNSS receiver has become an important tool for high-accuracy position services at present. However, in positioning measurement operations, characteristics such as long duration, bad conditions and diversified difficulties, bring serious effects on use efficiency and general applicability of GNSS receiver. Hence, measurement proposal

Foundation support: Young-aged Teachers' Basic Ability Improvement Project of Universities in Guangxi, No. 2017KY0629.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 71–77, 2020.
https://doi.org/10.1007/978-981-32-9244-4_8

which realizes "inertia-satellite" combined navigation based on integration of relevant inertia sensors has become a hot topic in the current research [2].

In current tilt survey researches [3], it usually using multiple inertial sensors [4] to measure carrier attitude angles (roll angle, pitch angle and azimuth). Then, using the rotation relationship of spatial attitude vectors to achieve the tilt compensation of geographic coordinates. The general applicability of this proposal is always restrained due to failure of normal work of the magnetic compass [5]. But till now, the proposals of reducing error factors, reducing cost and enhancing general applicability of tilt survey on the premise of ensuring measurement accuracy are rarely reported.

The paper analyzes relevant problems of the current mainstream tilt survey proposal. The research proposes an optimized tilt survey algorithm, namely High precision GNSS receiver tilt survey algorithm based on accelerometer. The algorithm enhances stability and universality of the attitude tilt survey. Finally, through comparative verification of measured data, it is verified that the thesis is better and more generally applicable than the current mainstream tilt survey algorithms.

2 Work Principles of Tilt Survey Based on Accelerometer

If the accelerometer is moving linearly at a constant speed or stays still, the triaxial output values of accelerometer can be recorded as Acc_X, Acc_Y and Acc_Z, which is described the force situation for CCS (carrier coordinate system). It's recorded as vector F^b. At the same moment, it only bears the gravity for GCS (geographic coordinate system) now. It is recorded as vector G^n, then:

$$F^b = [Acc_X \ Acc_Y \ Acc_Z]^T, \ G^n = [0 \quad 0 \quad 1]^T \tag{1}$$

According to the Euler's rotation theorem, it is used Matrix C_b^n describes a rotation matrix from F^b to coordinate system G^n. Its math connotation is that two coordinate system sharing the same origin point can coincide completely finally through a series of ordered rotation of fixed axis included angle. As follow:

$$G^n = C_b^n * F^b \tag{2}$$

Define θ, β and γ as roll, heading and pitch in the attitude angle, according to the Z-Y-X rotation principle in the aerospace field, it can be expressed as follows:

$$C_b^n = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\gamma & 0 & -\sin\gamma \\ 0 & 1 & 0 \\ \sin\gamma & 0 & \cos\gamma \end{bmatrix} \begin{bmatrix} \cos\beta & \sin\beta & 0 \\ -\sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

According to the formula (1), (2) and (3), relations between the roll angle θ and the pitch angle γ and the accelerometer output values can be solved:

$$\begin{cases} \theta = \arctan\left(\frac{\mathrm{Acc_y}}{\mathrm{Acc_z}}\right),\ \theta \neq 90 \\ \gamma = \arctan(-\mathrm{Acc_x}/\sqrt{\mathrm{Acc_y^2} + \mathrm{Acc_z^2}}) \end{cases} \tag{4}$$

The tilt angles indicate intersection angles between Inclined oxy plane and the geographical plane. It's defined that H1 is made pass point O. The H2 parallel to the H1 is used to cut oxy. Finally generated the geometric model shown in Fig. 1:

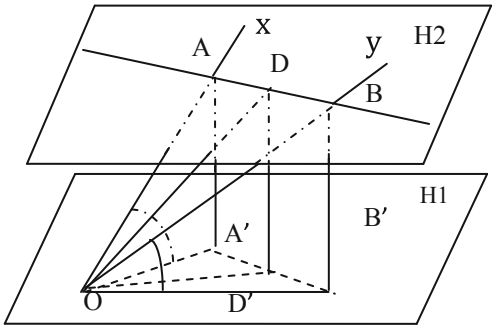


Fig. 1. Geometric model of carrier inclination

OA and OB indicate x axis and y axis of carrier coordinate system. D is the foot point on the AB edge, which passes O. Hence, the area of right-angled triangle AOB can be obtained through equivalent area conversion with two computation methods:

$$S_{\Delta AOB} = \frac{1}{2} * AB * OD = \frac{1}{2} * OA * OB \tag{5}$$

In right-angled triangle ODD′ AOA′, BOB′ and AOB, based on given definition, Horizontal tilt angle(δ) can be obtained as following relation:

$$\delta = \arcsin\left(\sqrt{\sin\theta^2 + \sin\gamma^2}\right) \tag{6}$$

When the tilt angle of the receiver can be obtained, combined with the GNSS receiver measurement parameters, tilt survey can be described as following Fig. 2:

Hence, attitude A, B and Z describe relations of projections of observation point of GNSS receiver under one attitude, with attitude A as the example, In the right-angled triangle OAA′, OA = H, ∠OAA′ = δ, the projection circle with known point as the circle center and OA' as the radius must pass the to-be-measured O. Hence, when randomly collecting three sets of projection circles, the equation set of circles can be listed and the general intersection solution can be obtained. In this way, the to-be-measured point can be computed.

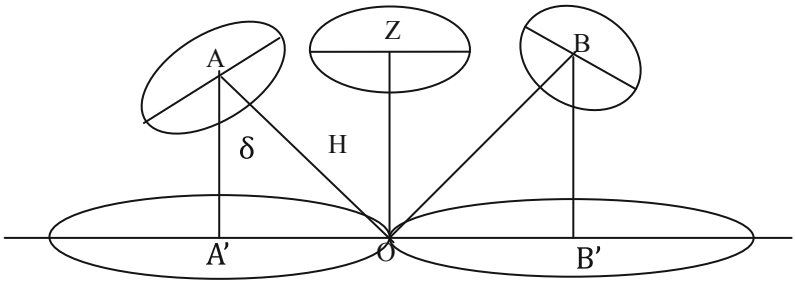


Fig. 2. Tilt survey geometric model based on acceleration sensor

3 Algorithm of Tilt Survey Based on Accelerometer

At first, the predefined array T_c stores and collects projection circles computed with tilt angle and positioning characteristics of the receiver under any attitude. It's defined that the data Circle describes the projection circle, involving serial number Index, center coordinates (X, Y), radius R and the geographic elevation H. Through actual Operations of the GNSS receiver, projection circles under tilt attitude are collected, computed and recorded to T_c for subsequent tilt survey computation.

When the acquisition completes the projection circle, According to T_c and in combination with the plane circle equation set, intersection points between every two circles are solved. Finally, according to intersection points that at least one of them share the similar solution, Extracting the feature intersection points that satisfy the condition by using the extraction algorithm. The specific process is as follows:

Input: T_c which stores three sets of projection circles.

Output: Values of geographic coordinates of to-be-measured point O.

Step 1. Traverse array T_c. According to the algorithm of solving intersection point with two circles, six intersection values between three circles are computed, it's Intersected in Pt1 and Pt2 of circles A and B. Intersection points of circles A and C are Pt3 and Pt4. Intersection points between circles B and C are Pt5 and Pt6.

Step 2. Extract Pt1 and Pt2. Work out their distances with Pt3, Pt4, Pt5 and Pt6, which are recorded as L13, L14, L15, L16, L23, L24, L25 and L26.

Step 3. Select the minimum value Min1 among L13, L14, L15, L16, Select the minimum value Min2 among L23, L24, L25, L26, Min1 and Min2 are compared. Step 4 will be started if Min1<Min2. Step 5 will be started if Min1>Min2.

Step 4. Recorded Pt1 as result. If L13< L14, recorded Pt3 as result, else recorded Pt4 as result. If L15< L16, recorded Pt5 as result, else recorded Pt6 as result.

Step 5. Recorded Pt2 as result. If L23< L24, recorded Pt3 as result, else recorded Pt4 as result. If L25< L26, recorded Pt5 as result, else recorded Pt6 as result.

Step 6. Output the Average data as to-be-measured point O.

The algorithm flow chart is as follows Fig. 3:

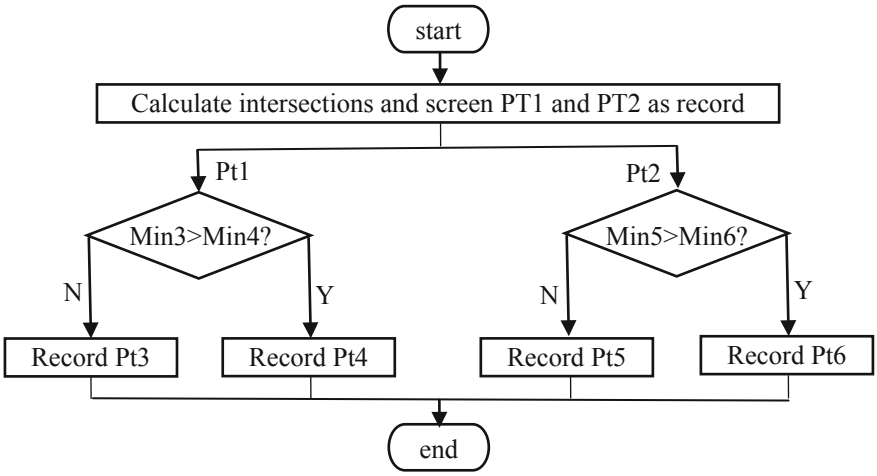


Fig. 3. The algorithm flow chart

4 Experiment and Result Analysis

When the GNSS receiver is operating with a tilt survey algorithm, Acquisition coordinate accuracy fluctuates with the superiority of the algorithm. Therefore, Accuracy of coordinates collected can directly reflect relevant performance of the algorithm. In order to verify the superiority of the algorithm in this paper, select a representative work area for testing: 1. Roofs with serious magnetic interference; 2. flower bed edges. And 50 groups of coordinates were collected with the Android measurement software in the above environment, as follow (Figs. 4, 5 and 6):

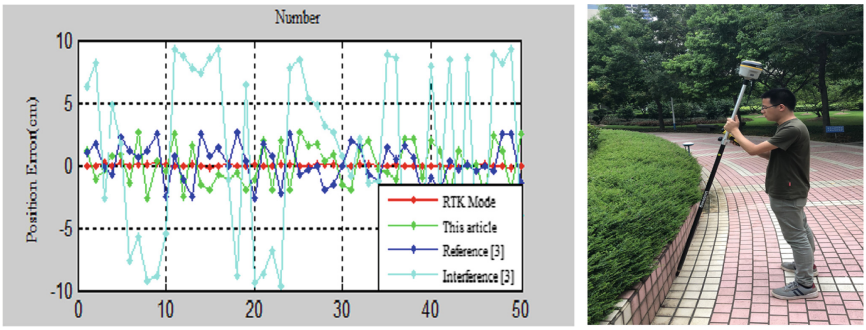


Fig. 4. Comparison of north coordinate and flower bed edges

Based on traditional GNSS Receiver measurement of millimeter accuracy result data, experimental results in above diagram indicate that the average error generated by the algorithm in the thesis aiming at geographic coordinates collected at any

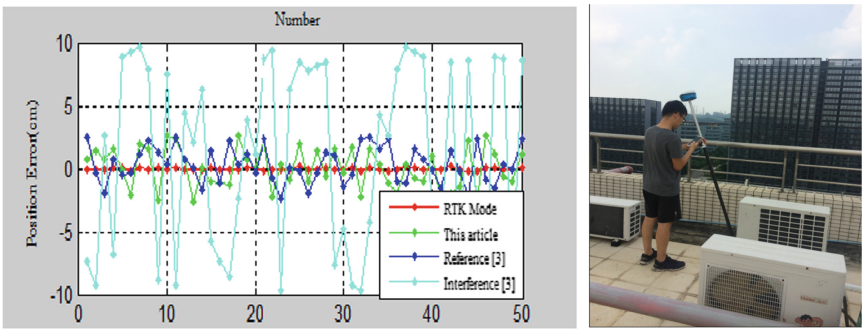


Fig. 5. Comparison of east coordinate and magnetic interference environment

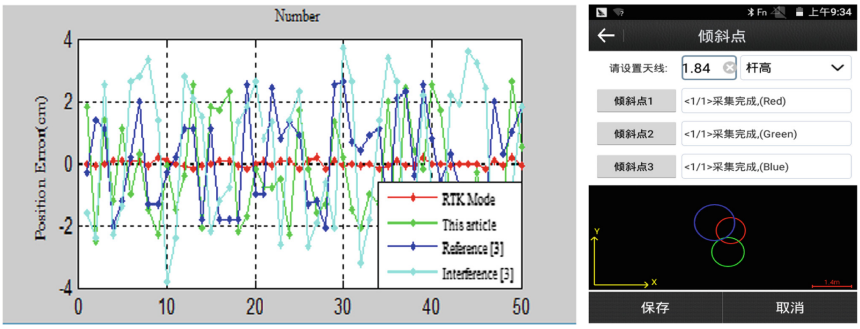


Fig. 6. Comparison of elevation coordinate and software

environment are stabilized within 3 cm. Through a series of calibration, the average errors of geographic coordinates calculated by the [3] algorithm also can realize the positioning accuracy within 3 cm. However, when serious magnetic interference exists in the environment, the average error for north and east coordinates and elevations obtained after tilt correction is larger than 10 cm and quite unstable. In general, the correction algorithm is meaningless under this case.

5 Conclusion

Aiming at the limit of high-precision GNSS receiver that use errors are large when mainstream tilt survey proposals are used during precision position services under strong magnetic interference and poor magnetic field environment, the paper puts forward a tilt survey solution proposal which only requires an accelerometer. On the premise of not changing the hardware of GNSS receiver, the proposal researches correlated geometric relations between carrier tilt angle and tilt pose during positioning measurement, constitutes equation sets with multiple sets of poses and solves general intersection solutions and works out coordinates of a to-be-measured point based on screening of characteristics to-be-measured points. Finally, accuracy analysis results of

measured data indicate that the positioning accuracy of algorithm in the paper can be stabilized within 3 cm. The positioning accuracy of current mainstream tilt survey algorithm is poor due to changes of magnetic field environment. It is verified that the algorithm proposed by the paper has better stability and general application performance. Meanwhile, the algorithm eliminates use of the magnetic sensor, so the measurable economic cost can be reduced for industrial generalization of pose tilt survey of the receiver. It is verified that the research of this paper has outstanding practical values in the precision position service field.

References

1. Liu, H.: Overview and latest development of global satellite navigation systems. In: 9th China Satellite Navigation Conference on Navigation Terminal Technology, China Conference, Harbin, pp. 124–128 (2018)
2. Teunissen, P.J.G., Khodabandeh, A.: Review and principles of PPP-RTK methods. J. Geodesy **89**(03), 217–240 (2014)
3. Tu, R., Shen, X.F.: Research on a new tilt survey algorithm based on RTK_GNSS receiver. In: 7th China Satellite Navigation Conference on Multi-source Fusion Navigation Technology, China Conference, Changsha, pp. 15–169 (2016)
4. No, H., Cho, A., Kee, C.: Attitude estimation method for small UAV under accelerative environment. J. GPS Solutions **19**(3), 343–355 (2015)
5. Li, M., Rouf, V.T., Thompson, M.J., Wen, J.K., et al.: Three-axis Lorentz-force magnetic sensor for electronic compass applications. J. Microelectromech. Syst. **21**(04), 1002–1010 (2012)

Rectangular Seal Location Algorithm Based on Digital Linear Feature

Liu Zhihui(✉) and Zhang Li

College of Software, Shanxi Agricultural University,
No. 1 Mingxian South Road, Taigu County, Jinzhong 030801, Shanxi, China
liuzhihui24@126.com

Abstract. The computer can reduce the interference of useless information by limiting the operation area and extracting the image of a single seal before identifying the authenticity of seal. Seal positioning is to determine the position of the seal in the image according to the geometric features of the seal. The rectangular seal region is extracted from the image to obtain a single seal image. The positioning of the seal image can greatly reduce the computation of the subsequent seal recognition and improve the accuracy and speed of the recognition operation. The difficulty of automatic locating rectangular seal lies in calculating center, side length and deflection Angle. This paper presents a method of automatic positioning of rectangular seal based on shape feature. Firstly, the color image of the seal is preprocessed to remove the useless information in the color image, and then the line segment element is used as the minimum unit to construct the line according to the rule of digital line. Quickly locate the rectangle in the image by comparing the length and direction of the line. The experimental results show that the positioning speed of the rectangular seal is improved while the accuracy is guaranteed. The position of the seal in the image can be quickly and accurately located.

Keywords: Seal positioning · Rectangle detection · Digital line

1 Introduction

Seals are used in many fields. Artificial multi-angle folding method can identify the true and false seal. However, this method requires a wealth of experience to judge, so the test results have a strong subjectivity [1, 2]. Using computer to quickly and accurately identify the seal has become an urgent problem to be solved. The operating object of seal authenticity recognition is a single seal image. The operating area is restricted before the seal is recognized by computer [3]. Locating the seal and extracting the image of a single seal can reduce the interference of useless information. Therefore, as a key step to judge the authenticity of a seal, automatic seal positioning is directly related to the speed and correctness of the judgment results. How to efficiently realize automatic seal positioning becomes an urgent problem to be solved. Rectangular seals are common in daily life. Gioi et al. [4] proposed a linear time segment detector, which can provide accurate results, controllable number of error detection, and does not require parameter adjustment. Lin et al. [5] have developed a rectangular detection algorithm based on line

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 78–84, 2020.
https://doi.org/10.1007/978-981-32-9244-4_9

detection. The idea is to find the line segment of a certain length range during line detection, which is used for the detection of aerial architectural images. Ren et al. [6] proposed a rectangular detection method based on gray projection integral extremum method. Liu [7] proposed a line segment detection algorithm based on global contour. This algorithm is a hierarchical method combined with Markov random field (MRF) model. On the one hand, the detection efficiency of these methods is directly dependent on the detection efficiency of straight lines; on the other hand, if the number of rectangles in the image is large, a large number of invalid line segment combinations will be generated, and the detection efficiency will be further reduced.

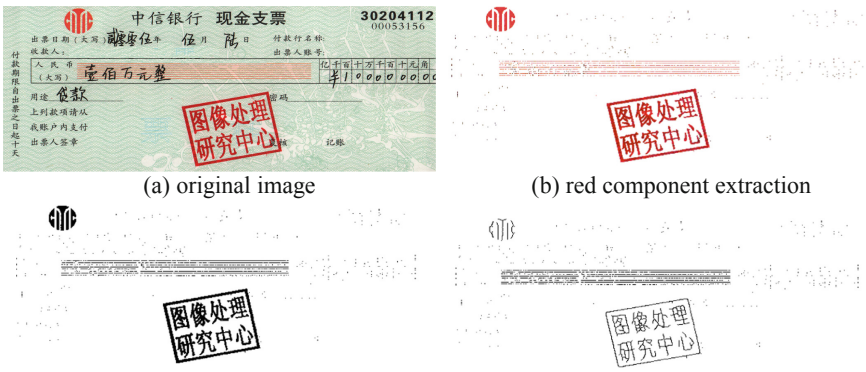
This paper proposes a rectangular seal positioning method. The goal of the algorithm is to complete the automatic positioning of the rectangular seal in the case of less user interventionl.

2 Research on Rectangular Seal Positioning Algorithm

In order to improve the speed of seal positioning and meet the real-time requirements, the grayscale image is first converted into a binary image. The skeleton diagram is obtained by thinning algorithm to highlight the shape characteristics of the seal [5]. Based on the detailed skeleton diagram, the method of locating seal is proposed. We detect the straight line according to the criterion of the straight line and the characteristic of the straight line. The line segment element is used as the minimum unit to construct a straight line. After detecting all the lines, the rectangles in the image can be quickly located by comparing the length and direction of the lines [6].

2.1 Image Preprocessing

In the collected test image, the seal region, background and text overlap each other. Mainly includes the red seal image and code number, gray or white background, black text. Although the color image contains abundant image information, it takes up a lot of space and increases the computation. Our aim is to retain the seal image as much as possible and remove the interference information.



(a) original image (b) red component extraction

(c) binary image (d) image refinement

Fig. 1. Results of preprocessing

The red component is firstly extracted based on the HSI color space model and the color image is grayed and binarized [7]. After binarization, the pixel value of the target object is 1 and the pixel value of the background is 0. Then, based on the refinement algorithm, a refined contour image is obtained from the binary image. Reduces a line with a certain width in the input image to a line with a width of 1 pixel. The topology of the original image is preserved while the information is reduced. Figure 1(a) is the original image, Fig. 1(b) is the red component extraction result graph, Fig. 1(c) is the binary image, and Fig. 1(d) is the refined image.

2.2 Basic Idea of Rectangle Positioning Algorithm

The image boundary uses 8 neighborhood to represent the chain code value [8]. Based on the chain code characteristics, it can be known that a digital line is composed of line segment elements. Based on this feature, the binary refinement graph was scanned, and all line segments in the graph were searched according to the definition of line segments. The deflection Angle between two adjacent segment elements is calculated. According to the included Angle, judge whether the two line segments are similar or not. If they are similar, connect the two line segments. The determination of whether the current segment element belongs to the line is based on the deflection Angle between the segment element and the line. A line segment element is used as the basic unit to construct a straight line. Thus the line in the image can be detected efficiently and accurately. Quickly locate the rectangle in the image by comparing the length and direction of the line. Find the center, length, width and deflection Angle of the rectangle.

The input of this algorithm is a refined binary image. The output is the center, length and width of the rectangle and the deflection Angle. The specific implementation steps of locating the rectangular seal are as follows:

Step1: we scan the refined binary graph to find the starting point $P_i(i = 0)$ of boundary tracking. P_i is the first pixel in the segment cell line_cell.

Step2: $i = i+1$, we track clockwise to find the next point P_i. If you return to the starting point of tracking or all points have been traversed, Step4; Otherwise, it determines whether P_i belongs to the current segment element line_cell. If it is, save P_i to line_cell and go to Step2; Otherwise, Step3 is performed for the current segment cell line_cell as a complete segment cell.

Step3: if the wireless segment element in the current line segment line_segment, add line_cell as the first line segment element of line_segment to line_segment; otherwise, judge whether line_cell belongs to line_segment. Let the number of pixels in the line_segment be N. Its starting coordinate is (X_{LSS}, Y_{LSS}). The endpoint coordinate is (X_{LSE}, Y_{LSE}). The number of pixels in the segment cell line_cell

is M. Its starting coordinate is (X_{LCS}, Y_{LCS}). The endpoint coordinate is (X_{LCE}, Y_{LCE}). Line segment element and line segment deflection Angle are shown in formula (1):

$$\begin{cases} \theta_{LC} = \arctan\frac{Y_{LCE}-Y_{LSS}}{X_{LCE}-X_{LSS}} \\ \theta_{LS} = \arctan\frac{Y_{LSE}-Y_{LSS}}{X_{LSE}-X_{LSS}} \end{cases} \quad (1)$$

If formula $|\theta_{LC} - \theta_{LS}| \leq {}^{\alpha}/_{M+N}$ is satisfied, line_cell is considered to be a part of line_segment. If not, treat the line_segment as a complete line segment. Store the segment and empty the line_segment, where A is α threshold. Initialize the segment cell line_cell with point P_i. Similarly, initialize the line_segment with the current segment element line_cell and go back to Step2.

Step4: if the deflection Angle of two line segments is close and their endpoints are adjacent, then merge the two line segments.

Step5: successively traverse the straight line segment with the number of pixels greater than the threshold value T_{seg}. So let's see if that satisfies formula $\Delta\theta = |\theta_i - \theta_j| < T_\theta$. T_θ is the Angle threshold. Put each pair of parallel lines into the structure;

Step6: traverse each pair of parallel lines successively. So let's see if that satisfies formula $\big||\alpha_i - \alpha_j| - 90^\circ\big| < T_\alpha$. T_α is the Angle threshold. The four sides of the rectangle are obtained.

Step7: eliminate the pseudo-rectangle. If only the length and direction of the line are used to determine whether it is a rectangle, then it is inevitable that the dotted area in Fig. 2 will be treated as a rectangle. So we need to get rid of the pseudo-rectangle. We compute the rectangle vertices P_1, P_2, P_3, and P_4 from each side of the rectangle. The number of pixels on the straight line segment P_1P_2, P_3P_4, P_2P_3 and P_1P_4 was counted and compared with the length of each side. Only when condition (2) is met, the rectangle is judged to be true.

$$|C_{12} - \xi_1| < T_\xi\xi_1, \quad |C_{23} - \xi_2| < T_\xi\xi_2, \quad |C_{34} - \xi_1| < T_\xi\xi_1, \quad |C_{41} - \xi_2| < T_\xi\xi_2 \quad (2)$$

Where, C_{12}, C_{23}, C_{34} and C_{41} represent the number of pixels near the straight line segment of P_1P_2, P_2P_3, P_3P_4 and P_4P_1 respectively. ξ_1 and ξ_2 represent the length of the rectangle, and T_ξ = 0.2 is the threshold.

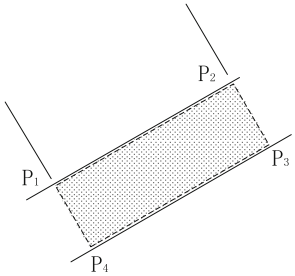


Fig. 2. Pseudo-rectangle

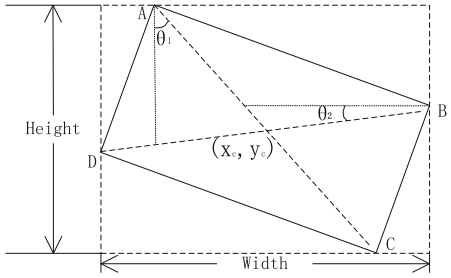


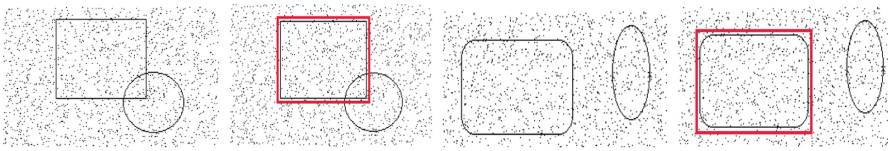
Fig. 3. Schematic diagram of rectangle parameters

Following the above steps, the rectangle in the image can be detected. Figure out the center (x_c, y_c), side length and deflection Angle θ_1 and θ_2 of the rectangular seal, as shown in Fig. 3. That's where the rectangular seal is in the image. The pixel points centered on (x_c, y_c) and within the range of (x_c, y_c) are intercepted to extract the rectangular seal image, where $Height = AC \times \cos\theta_1$, $Width = BD \times \cos\theta_2$, $Height' = Height + \xi$, $Width' = Width + \xi$ and ξ takes the empirical value of 40.

3 Experimental Results and Analysis

In order to verify the correctness and validity of the algorithm, experiments were carried out with Matlab on a computer with Intel 3.39 GHz and 4 GB memory. We mainly analyze the accuracy of the algorithm and the ability to resist noise interference. The simulation images and real life images were tested.

Experiment 1: Simulation Graphic Detection Example
In order to verify and compare the detection performance of the algorithm, the following two groups of experiments were conducted, as shown in Fig. 4. Two simulated images with salt and pepper noise were tested respectively. In Fig. 4(a), the ellipse and rectangle intersect, and in Fig. 4(c), the ellipse and rectangle are separated. Table 1 shows the comparison of detection performance between the algorithm in Reference [7] and the algorithm proposed in this paper.



(a) test picture1 (b) result of test picture1 (c) test picture2 (d) result of test picture2

Fig. 4. Simulation test picture

Table 1. Comparison table of simulation image detection performance

Noise ratio	Testing time		Missing rate	
	Reference [7]	Algorithm in this paper	Reference [7]	Algorithm in this paper
0%	1.05 s	0.64 s	0%	0%
5%	1.23 s	0.9 s	0%	0%
10%	1.38 s	1.15 s	6%	0%
20%	2.7 s	1.69 s	13%	8%
40%	4.63 s	2.01 s	38%	12%
70%	9.08 s	4.38 s	73%	56%

Experiment 2: Real Image Detection in Real Life
Due to the different seal test map, the depth is not the same, even for the same seal test map, the brightness of different parts may also be different. In this paper, different seal pictures were tested respectively. In this experiment, 50 seal images were selected as experimental materials. All images were downloaded from the Internet. Figure 5 shows the detection results of real images. The results show that the algorithm can accurately detect the location of the rectangle, it has a certain anti-noise ability, can detect the complex image in the rectangle.

Table 2. Comparison table of real seal image detection performance in life

Algorithm	Average running time	Accuracy	Total number to be tested	Amount of residual
Reference [7]	2.16 s	87%	50	7
This paper	1.23 s	95%	50	3

Table 2 shows the comparison of detection performance between the algorithm in literature 7 and the algorithm in this paper. By analyzing the experimental results, three cases of seal omission were summarized. First, the color of the seal was seriously degraded, and some areas were covered by the background color. Secondly, the seal area was divided into several small areas due to the screen creases and other screen contents. Third, the seals are of special shapes, such as gourds.

(a) original picture (b) positioning result of rectangular seal

Fig. 5. Invoice

4 Conclusion

This paper proposes a method of automatic positioning of rectangular seal. The algorithm can effectively extract the red component of the test image and locate the specific position of the rectangular seal on the gray image. Reducing the area of the subsequent processing area is of great significance for improving the speed simplification algorithm. Experiments show that this method can detect the rectangular seal in the image accurately. Under the condition of noise and incomplete rectangle, the detection result is better. It lays a good foundation for seal authenticity decision and has certain application value. The future work focuses on how to improve the detection accuracy of seal area and the positioning of special shape seals.

References

1. Wei, Z., Wei, M.: Legal risk prevention in corporate seal management. Friends Secr. **4**(7), 11–12 (2013)
2. Zhao, Y.: Analysis on the standardized management of corporate seals. Mall Mod. **6**(26), 180–182 (2012)
3. Xie, Q.: Known angle straight line detection method based on template filtering. Manuf. Autom. **34**(12), 83–84, 87 (2012)
4. Grompone von Gioi, R., Jakubowicz, J., Morel, J.M., Randall, G.: A fast line segment detector with a false detection control. IEEE Trans. Pattern Anal. Mach. Intell. **32**(4), 722–732 (2010)
5. Lin, C., Nevatia, R.: Building detection and description from a single intensity image. Comput. Vis. Image Underst. **72**(2), 101–121 (1998)
6. Ren, L., Liu, K., Zhang, H., et al.: Rectangular detection of the extreme value method of gray projection integral. Comput. Eng. **38**(8), 159–163 (2012)
7. Liu, Y., Ikenaga, T., Goto, S.: A novel approach of rectangular shape object detection in color images based on an MRF model. In: IEEE International Conference on Cognitive Informatics, vol. 1, no. 11, pp. 386–393 (2006)
8. Shang, Z., Liu, M.: Straight line detection algorithm based on Freeman criterion. J. Comput. Aided Des. Graph. **17**(1), 49–53 (2005)

Big Data Based Decision-Making Support System Design for Efficient Analysis of the Performance of Software Education

Ji-Hoon Seo and Kil-Hong Joo(✉)

Gyeongin National University of Education,
115, Sammak-ro, Manan-gu, Anyang-si, Gyeonggi-do, Republic of Korea
sserz@naver.com, khjoo@ginue.ac.kr

Abstract. This study is aimed to analyze the performance of software education that has been implemented thus far centering on the opinion mining visualization analysis technique based on big data related to software education collected over the last five years such as data from portal site news, SNS service, Internet café, and other performance data and develop a big data based decision-making support system for Korean style software education in order to derive agendas that will be discussed hereafter. In addition, through the prediction of more advanced future-oriented education systems and analysis of performance factors based on the foregoing, solutions for big data based decision-making support prediction analysis can be prepared hereafter in the field of education and measures for cultivation of creative convergence talents and promotion of software education can be sought so that the directions of mid/long-term development of Korean style software education can be accurately set.

Keywords: EDM · Software education · Decision support system · Big data

1 Introduction

Recently, R&D project plans linked to the fourth industrial revolution on which the South Korean government focuses have been implemented. Those projects include national projects utilizing big data, which are applied to diverse fields such as ICT, BT, and Smart-City and are fully supported by the government including the establishment of new big data related organizations and laws. On the contrary, measures to utilize big data in the field of education are still slow and are hardly supported at the government level [1, 2]. However, thanks to the excellent IT infrastructures, projects utilizing big data are now gradually progressing in the field of education and since it has become possible to make massive data accumulated over several years into big data, such pieces of big data information can now be linked with each other for analysis [3, 4]. Meanwhile, since various departments have been implementing systematic programs intended to improve the constitution of education and cultivate talented persons such as software education and creativity, personality, and convergence education centered on learners' problem solving ability and computational thinking skills, diverse and huge data and pieces of information lie scattered [5, 6]. While analyzing South Korean software

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 85–90, 2020.
https://doi.org/10.1007/978-981-32-9244-4_10

education systems over the last five years, we have achieved eye-opening performance in reducing the sense of difference between education and the people. However, there has been no method to concretely quantify and measure the performance and although the results of individual analyses exist, there is no appropriate tool that can organically link those pieces of information with each other for analysis. In addition, in South Korea, still there is no system at all that can construct big data using unstructured texts to predict decision making for educational outcomes. Therefore, in this study, based on big data analyses, a decision-making support system that can seek the performance and the direction of future oriented development of software education that has been reorganized into compulsory education at elementary/middle schools from this year will be designed, the education policies implemented by various government departments will be evaluated in general, the big data based on the unstructured data accumulated in online media will be analyzed centering on opinion mining among the three unstructured data analysis techniques, that is, text mining, data mining, and opinion mining, and based on the foregoing, a big data based decision-making system design model that can derive the prediction of changing future education systems and the direction of promotion of polices as quantitative performance will be proposed.

2 Related Work

2.1 Core Strategies of South Korean Software Education

Focusing on the stable settlement of SW education in the South Korean school field, computing thinking ability was defined as 'the thinking ability to efficiently solve problems that may occur in students' daily lives based on the basic concept and principles of computing.' It has been pointed out that although the importance and necessity of SW education have been sufficiently recognized, for SW education to be effective, the goal and content of the curriculum should be concretized and the operation of the curriculum should be flexible. Accordingly, although computing thinking centered studies have continuously achieved good outcomes, such studies show a shortcoming of being not sufficiently realistic.

2.2 Operation Guidelines for South Korean Software Education

The SW education implemented in the 2009 revised curriculum comprises one unit of the subject manual training for the 5th grade of elementary school and an elective subject (information) for middle school, which is selectively taught at some schools. Whereas the content of education of the subject information in the existing curriculum is focused on the utilization of information and communication technologies, the content of education in the operation guidelines for software education has been said to be focused on the reinforcement of information ethics and the improvement of problem solving ability utilizing algorithms and programming (Ministry of Education - KERIS 2015). According to the 2015 revised curriculum, in the case of elementary school, the existing unit information centered on ICT utilization should be changed into a large unit centered on basic knowledge of SW to implement the education with a content, which is centered on program solving processes and experience in algorithms and programming and includes the fostering of information ethics consciousness.

3 Proposed Method

For the big data based decision-making support system proposed in this paper, a dictionary context based pre-treatment process was implemented through which all software education related unstructured data distributed in online media were collected and inappropriate keywords were removed, and a process was configured for design of an efficient decision-making support system for analysis of the performance of software education performance.

3.1 Development of Korean Style SW Education Data Collection, Storage, and Processing Technologies

In this study, to reduce the dependency on real-time data collection and present correct directivity through the prediction of the future of software education, the python-based crawling technique was used as a method to collect raw data. Atypical data such as online news articles, blogs, social networks (Twitter and Facebook) from 2012 when SW education became an issue to 2017 were collected and classification, preprocessing, definition for data standardization, and bidirectional sharing systems were grafted centering on big data storage and processing technologies for decision making support systems.

3.2 Development of a Data Integration System for Prediction and Analysis of Atypical Data for SW Education

Since configuration systems that are stable in and suitable for physical environments such as DW (Data Warehouse) for decision making analysis, an environment where atypical data can be distributed for processing utilizing storage servers and master servers in constructing big data was constructed. An accuracy improving technique was designed to extract reliable affirmations and negations from text sentences in creative education contents using opinion sentiment dictionaries trained on individual atypical text documents. For ETL-based data extraction, transformation, and loading, a disk sharing mode system was constructed to store and refine past raw data, database modeling was constructed utilizing RDBMS, and Map Reduce-based and Stream-based big data environments were designed (Fig. 1).

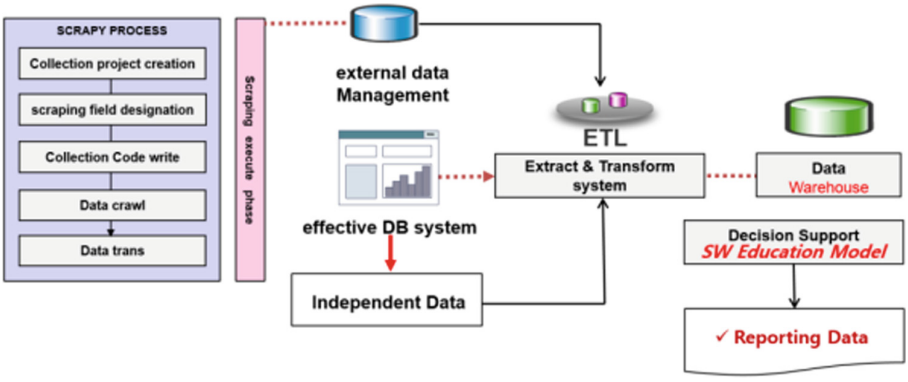


Fig. 1. Software education big data decision making support and storage model

3.3 Korean Grammar Based Pre-treatment Design

In this study, outliers, missing values, or wrong values are not shown in the data values because the preprocessing was performed based on the collected unstructured data. In particular, filtering was conducted to extract highly important words in sentences or important words that could be sentiment words. In the United States, in cases where sentiment words for English grammar are extracted, the words can be easily preprocessed utilizing the SWN (SentiWordNet). A characteristic of the SWN is that it contains synonym sets based on 147,278 English words belonging to those parts of speech such as nouns, verbs, adjectives, and adverbs and that it is stipulated that the sum of the sentiment values of affirmation, negation, and neutrality in each set should be 1. That is, the SWN is specialized for extraction of sentiment words from sentences in English grammar because of the structure of its grammar algorithm and the structure makes it difficult to sentiment words from Korean sentences and becomes a factor to degrade prediction accuracy. Therefore, in this study, the SWN was excluded and to efficiently derive sentiment words from sentences in Korean grammar, the rules as shown in Table 1 were applied to perform word filtering.

Table 1. Word stemming filter process.

Filtering rules applied for the construction of a sentiment dictionary
1. Remove special characters, English words, and disused vocabularies
2. Remove meaningless terms and one-letter texts
3. Separate the same words in the form of a conjunction to classify the natures of sentiment words
4. Distinguish homonyms and synonyms from each other
5. In the case of abbreviations and newly coined words, reflect only those that have been registered in WiKipedia or a Korean language dictionary

4 Experimental Result

The In this study, to analyze the performance of Korean style software education, massive data were stored and accumulated and related preprocessing was performed. In addition, to improve the accuracy of the analysis, the decision-making support design technique for Korean grammar was used and the resultant antonym correspondence, homonym and synonym rule classification, and associated word decomposition were performed. Thereafter, the existing results of analysis of texts for software education and the results of reputation analysis conducted through the decision-making support system designed in this study were compared for accuracy.

The environment of comparative analysis is as follows. First, in order to increase the accuracy of software education related text words, 41,000 words that appeared regularly at high frequencies were extracted by applying the related rules, and some of the nouns maintaining a close relation with the positive units among the extracted words were included in the important words. Second, all vocabularies recognized as

one letter were excluded in the filtering process, and verbs and antonyms with high frequencies of association between words were extracted. With regard to the performance analysis, the results of reputation analysis using the existing sentiment words were compared with the results of reputation analysis using the design technique presented in this study (Fig. 2).

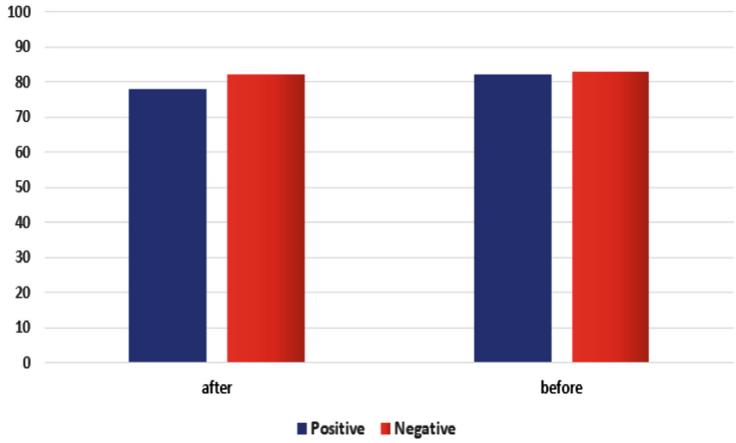


Fig. 2. Performance analysis

According to the results, the accuracy of the existing technique was shown to be 78% in the case of positive word and 82% in the case of negative words. The accuracy of the negative word was 78% Accuracy was 82% accuracy. The accuracy of reputation analysis not lower than 80% corresponds to stable values. The accuracy of reputation analysis that introduced the design technique presented in this study was shown to be 82% in the case of positive words with an improvement by 4% compared to the existing model and 83% in the case of negative words with an improvement by 1% compared to the existing model. Although the results did not show very high performance with 1% improvement from sentiment analysis, the proposed technique is expected to show higher accuracy when the data for word analysis are larger.

5 Conclusion

In this study, a decision making support platform was designed for analysis of the performance of Korean style software education. If the proposed big data analysis platform is used, change factors for diverse education environments can be analyzed and future-oriented agendas for creative talent cultivation and coding education can be discovered. Learner education systems can be improved so that excellent creativity with good problem solving ability to solutions based on new and unique perspectives can be discovered and present ways to suggest diverse solutions for similar problems and if the data in the field of education that have been already accumulated,

ways to implement education suitable for students can be presented. In addition, to compare with previous analysis methods, whereas methods such as brainstorming, Delphi, and expert panels were used in the past, this study enabled the presentation of data based analysis results through big data decision making support platform using research literature, media articles, and related reports. Therefore, the results of this study are expected to be widely utilizable in the field of convergence in the domain of humanities too.

Acknowledgments. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2017R1D1A1B03029292) and the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF- 2017S1A5A2A01025431).

References

1. Chen, M., Mao, S., Liu, Y.: Big data: a survey. Mob. Netw. Appl. **19**, 171–209 (2014)
2. Wu, X., Zhu, X., Wu, G.-Q., Ding, W.: Data mining with big data. IEEE Trans. Knowl. Data Eng. **26**(1), 97–107 (2014)
3. Tang, C., Liu, C.: Method of Chinese grammar rules automatically access based on association rules, In: Proceedings of the Computer Science and Computational Technology, vol. 1, pp. 265–268, 20–22 December. ISCSCT, Shanghai (2008)
4. Khan, I.A., Choi, J.T.: An application of educational data mining (EDM) technique for scholarship prediction. Int. J. Softw. Eng. Appl. **8**(12), 31–42 (2014)
5. Xu, Y., Li, Y., Shaw, G.: Reliable representations for association rules. Data Knowl. Eng. **70** (6), 555–575 (2011)
6. Klein, D., Tran-Gia, P., Hartmann, M.: Big data. Informatik-Spektrum **36**, 319–323 (2013)

Research Process Traceability Based the RD Relationship

Jeong Ah Kim[1(✉)], SunTae Kim[2], JaeYoung Choi[3], Ji Young Lee[3], and Young Hwa Cho[3]

[1] Department of Computer Education, Catholic Kwandong University, 24, Beomil-ro 579 beon-gil, Gangneung-si, Gangwon-do, Republic of Korea
clara@cku.ac.kr

[2] Department of Software Engineering, CAIIT, Chonbuk National University, 567 Baekje-daero, Deokjin-gu, Jeonju-si, Jeollabuk-do, Republic of Korea
stkim@jbnu.ac.kr

[3] Department of Computer Engineering, SungKyunKwan University, 2066 Seobu-ro, Jangan-gu, Suwon, Gyeonggi-do, Republic of Korea
{jaeychoi,jiyoung,hoyh2285}@skku.edu

Abstract. It is not easy to define the well-defined process for software research project but it is important to keep the tracking the progress of the projects. Without the process, it is not easy to measure the progress since we cannot define the unit of measure. In this paper we suggested RD(research descriptor) concept to define the unit of measure. Also we suggested RD relation to trace the progress the research workflow. With suggested mechanism, we can monitor the research project process and visualization the progress.

Keywords: Software research · R&D project · Traceability · Deep learning algorithm

1 Introduction

At a time when highly reliable software (SW) engineering is improving the information technology (IT) field, it is essential to reduce maintenance costs by securing original technologies for preemptive SW quality control and auditing to support transparent and efficient development. Additionally, when the research scope and goals established at the initial stage change during development, because of uncertainty and variability of the research and development (R&D) field, a constant-monitoring system, based on research contents, using semantic analysis to support the logical and systematic reflection process, is required. A research descriptor (RD), used as a meta-model to express matters (e.g., research, process, indicators, and guidelines) occurring during the R&D process, should be developed to support multiple platforms so that the user can freely configure and describe the format depending on various work environments and preferences, such as SW development stages, individual research and development styles, research contents to be described, and tools to be used during creation (e.g., Hangul, Microsoft Word, Microsoft PowerPoint). RDs are data used to express work products of the research process and final results of research. In this study, we propose

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 91–97, 2020.
https://doi.org/10.1007/978-981-32-9244-4_11

a method to define the traceability between research processes and work products based on semantics by defining the RD relation, and a method of verifying its validity through machine learning. We also propose a visualization method that facilitates monitoring the RD relation.

2 RD Definition

RD has the following properties that serve as requirements when defining its structure [1].

(1) **R&D process and work product information management**: RD should express and store information on the process in a series of procedures for conducting R&D projects and information development on various types of work products produced by each process.
(2) **Traceability support between R&D processes, between the process and the work product, and between the work product and the work product:** RD must include traceability information between RDs. Here, RD includes management RDs (mRD) storing process information and engineering RDs (eRD) storing work-product information.
(3) **Consideration of flexible/free style R&D process and work product information management**: RD should manage not only fixed processes and work products, but also flexible/free style processes and work products in the future.
(4) **Used for testing and as a unit of search/reuse:** RD containing processes and work product information can be used for testing and for reuse.
(5) **Support of hierarchical structure, support of structures containing other RDs**: The inside of a work product is defined by using the hierarchical relation. Thus, there are multiple sections inside one document, and there are multiple subsections in one section. Therefore, RD should support the definition of this hierarchical structure.
(6) **Support of process quality (monitoring indicator)**: Because a single RD has a relation (traceability) with multiple RDs, it is possible to measure their quality. RD should be able to check their quality through a monitoring indicator.

An SW R&D project can be divided into a process and a work product. The process area comprises project–phase–activity–task, and each process has attributes, such as title, period, purpose, and participant. The process area can have more components and attributes or simpler components and attributes to meet R&D properties. However, it is defined in this manner to demonstrate RD requirements. In an activity or task, various work products are created while conducting the corresponding process. A work product also comprises various work-product items. The properties (requirements) for RD include all contents of the work product and items stored as RDs, with information on each element (i.e., project, phase, activity, and task) in the R&D project. Additionally, the contents for traceability between processes (comprises), for traceability between the process and its work product (produces), and for traceability between work products (includes) are defined as traceabilities between RDs.

3 Definition and Verification of RD Relation Type (Traceability)

3.1 Definition of Traceability in the General SW Field

The definition of traceability is expressed quite differently, depending on research and industry. The definition of traceability is applied uniquely to each field via measurement standards of the mechanical industry, material engineering, software engineering, and food processing involved [2, 3]. The most standardized form of traceability is the association of uniquely identifiable entities containing attributes of time required to verify the relationships. Generally, the SW field defines traceability as the ability to manage the relationship between requirements and other work outputs [4, 5]. Verification and validation of the work products (e.g., design documents, various implementations, source code, executable files, and test case plans), derived from requirements, satisfy and fully reflect all functions described in requirement specifications as activities that can support the traceability defined above. However, it is difficult to conclude whether the generic definition of traceability can accommodate all aspects of SW development. Traceability should rather be defined not only as traceability between different work products but also as traceability that can be applied to various ranges, such as traceability for the inside of a specific work product or traceability for changes in a work product caused by changes over time.

3.2 Definition and Utilization of RD Relation Type (Traceability)

RD traceability is the relation between multiple RDs. The relation type is a format that establishes the relationship, setting a name that can describe its nature. Information on the RD relation type should be used not only to identify the relation between RDs and to infer additional traceability, but also to establish and verify traceability using semantic analysis. Until now, the unit of traceability or monitoring in research has been the unit of the document or model. With the recent focus on model-based development, research on the information model establishing traceability between models has advanced. For model-based traceability, because the transformation flow between models is already defined, trace connections can be created based on model transformation. Because this study conceives managing tasks based on RD, it cannot predefine the mapping between model elements created during the transformation flow between models or the model transformation process. Nonetheless, conducting this research creates various contents to achieve a research goal. Thus, there is a refined relationship between these contents. This study allows researchers to define the relation between RDs to identify the research flow, as it provides researchers with the degrees of freedom to define research contents as RDs or as objects of management. To define the relation between RDs, we proposed four basic relation types (Table 1 and Fig. 1).

Table 1. 4 relation type for defining the traceability

Relation type	Meaning	Utilization	Example
Progress	Meaning of the activity results related to the evolution of the research process, as relation from an existing RD to a new RD or a modified RD	Track the flow caused by changes by introducing new attempts, new perspectives, and new methods, or by finding and fixing errors	Ontology-based reasoning module-analogical reasoning module
Materialization	Expression of the relation between contents with low-level abstraction and contents with high-level abstraction	Relation between RDs being made, progressing from a conceptual level to an increasingly realizable level	Goal-use case Use-case data model; Component-class
Segmentation	Expression of hierarchical and compositional relations between contents	A relation dividing complex and large contents into small and manageable units	Goal- constraints; System-subsystem
Same	There is no change in content, yet the relation between RDs that describes the same content in different expressions	The management of relations between contents expressed from different perspectives on the same topic	Figure-description; Formula-description; Requirements-definition

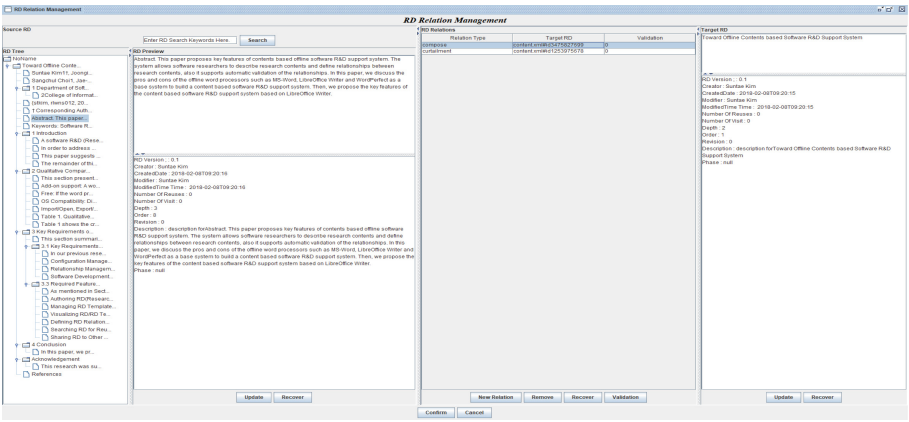


Fig. 1. RD relation management window.

We implemented a relation definition tool to define RD relation types and piloted it. The RD relation management tool first provided the RD structure of documents in a tree-format on the left side of the window to support the overall management of RD relations. It showed researchers a preview and RD meta information when an RD was

selected. With those, RDs related to the selected RD were also shown. If the user selected a related RD, information regarding the related RD was displayed on the screen in the same manner. RD relation management provided not only RD inquiries, but also set-up functions for adding new relations, deleting or modifying existing relations, and deciding the validity of RD relations.

3.3 Development of an RD Relation Recommendation Tool Based on Machine Learning

We implemented a semantic analysis verification tool to decide the validity of RD relations. This enables researchers to accumulate good semantic relation data by calculating the suitability of relations specified between RDs. This also provides document recommendations that researchers will need in the future. The algorithm for deciding the validity of the RD relation type is as follows.

(1) **Same**: there is no semantic or quantitative change between RD A and RD B
(2) **Materialization**: There is no semantic change between RD A and RD B, yet there is a quantitative increase in RD B
(3) **Segmentation**: Some of the content in RD A are deleted from RD B
(4) **Progress**: Content that does not exist in RD A is added to RD B

The semantic relation analysis algorithm follows:

```
step 1 Data preprocessing
       (e.g.,    sentence    classification,    dictionary
creation)
Step 2 Sentence vector (D-dimensional) modelling that can
        reflect   semantics   (e.g.,   word2vec,   paragraph
vector)
Step  3  Calculate  the  semantic  similarity  for  each
sentence
        pair between RD A and RD B
Step 4 For each sentence, measure the similarity of all
       sentences of the corresponding RD, and then select
       the maximum value
Step  5  Ensure  that  the  maximum  similarity  value  is
greater
       than a certain threshold for all sentences
Step 6 If all sentences in RD A and RD B are above
            the threshold,
       If the number of sentences of RD B is more than
            that of RD A, then do materialization
       Else, same
       Else
            If there is a sentence in RD A that does not
               exceed the threshold, then do segmentation
            If there is a sentence in RD B that does not
               exceed the threshold, then do progress
```

4 Monitoring Through the Visualization of RD and RD Semantic Relations

In this study, we effectively visualized the contents of information based on the semantic relation within or between RDs and provided an effective function of constant-monitoring by implementing a tool that can show the analysis results of life-cycle traceability from various angles. It is a system that applies this visualization method as one of information and knowledge and defines interconnection relationships between contents comprising one or different documents while creating new documents. To enable the flow, search, track, and mutual verification of contents by giving stereoscopic and dynamic meanings via connections, we aimed to show the relation diagram and contents having relationships through the interconnections between RD creations by implementing the visualization of the RD-centered relations that visualize them.

The main subjects of each RD, obtained through the semantic analysis and the subject to express flow, were collected. In the visualization, RD was expressed as a node graph, and the relation between RDs was expressed as a link, or a line, through which RD relations could be identified and tracked. Each RD should see information differently according to size, documents, groups, content, process of creation, type of corresponding document (e.g., technical report, work product, result report), or content selection. To reduce visual complexity in these situations, the same or similar subjects are combined on the graph, and new topics are separated. RD has an interrelation, and the flow of the relation becomes a process of R&D. Therefore, it is necessary to visualize to confirm the nature of a specific relationship via traceability or to identify the flow of the corresponding relation centered on a specific RD. The nodes in the graph are produced as work-product documents (i.e., RDs) using the RD authoring tool. The status of the RD can be visualized with a connection line having directivity, based on the interrelation stored, based on the established contents.

5 Conclusion

So far, traceability is the association of uniquely identifiable entities defined in well-defined template-based work products containing attributes of time required to verify the relationships. In this study, we suggested RD(research descriptor) and RD relations to make traceability. These concepts to make possible to define the traceability relationship between the entities which are not pre-defined. While the research is going, several RD can be produced to contain the any kinds of research. With RD relation type it make possible to identify the relation between RDs and to infer additional traceability. For verifying traceability, we developed the semantic analysis techniques based on deep learning mechanism. Also, visualization framework give the image of project progress.

Acknowledgment. This research was supported by Next-Generation Information Computing Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2014M3C4A7030503). JeongAh Kim is the corresponding author.

References

1. Cho, Y.H., et al.: Development of research content based real monitoring system using the content analysis. Technical Report (2016)
2. Paech, B., von Knethen, A.: A survey on tracing approaches in practice and research. Technical Report IESE Report Nr. 095.01/E, Fraunhofer - Institut Experimentelles Software Engineering (2002)
3. Palmer, J.D.: Traceability. In: Software Requirements Engineering, pp. 364–374 (1997)
4. Ramesh, B., Jarke, M.: Toward reference models for requirements traceability. IEEE Trans. Softw. Eng. **27**(1), 58–93 (2001)
5. Galvão, I., Goknil, A.: Survey of Traceability Approaches in Model-Driven Engineering

Improvement Design of Smart Toilet Interactions Based on Peak-End Rule

Yujia Wu and Bowen Sun(✉)

Beijing Institute of Technology,
No. 5, Zhongguancun South Street, Haidian District, Beijing, China
3120181759@bit.edu.cn, doersun@gmail.com

Abstract. Technology is recognized as a vital measure to promote the transformation and the upgrading of medical industry when health management has entered the intelligent era. Smart toilets, considered to have great potential for monitoring health information at home, emerge in our daily life. This paper is conducted to ameliorate the key interactions between human and smart toilet and based on peak-end rule, effectively improving user experience. A new design approach is applied to finds the peak and end moments in using process, and finally identifies four key interactions through the method combining user interview and Kano model. This paper details how the application of the new approach improves the interactions between human and smart toilet, contributing to a smoother product operation process which creates a better user experience.

Keywords: Interaction · Smart toilet · Peak-end rule · User experience

1 Introduction

Development of information age integrates the Internet of Things into people's modern life. Facing the acceleration of China's aging population, demands on intellectual product to make people do health checks more conveniently are mounting. The emergence of medical products embodied with intelligence, such as smart blood pressure meters, symbolizes that interaction between human and computers is more frequent and essential [1].

In decades, user experience has evolved into a core concept of human-computer interaction. It has inflamed a heated discussion among practitioners and researchers from various disciplines who look for ways to expand their understanding of "pleasant experience" [2–4]. Kujala et al. [5] said that the goal of user experience design in industry is to improve user satisfaction and loyalty through the utility, ease of use, and pleasure provided in the interaction with a product. Some researches on the design of smart toilets have already been done. Xu et al. [6] have used the QFD and FBS models to develop the versatility of toilet, which can accurately and effectively convert user requirements into design requirements. Nevertheless, as an innovative intelligent product, the research on user experience of smart toilet in domestic and foreign academic circles is few and limited and relative research tends to focus on technology, perfection or user needs of vulnerable population [7–9]. Therefore, combining the theory in other

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 98–103, 2020.
https://doi.org/10.1007/978-981-32-9244-4_12

disciplines brings new opportunities for improving user experience of smart toilet. Peak-end rule, known as a psychological theory, has also been used by Do et al. [10] to improve and evaluate designs.

Combining Peak-end rule theory, Kano model and user interview, this paper aims to apply cross-disciplinary design approach to smart object. Peak-end rule claims that people's memory of experience is determined by the feelings at the peak and the end moments. Finding out the "peak" and "end" moments of the smart toilet and carry out corresponding design improvement can not only enhance user experience, but also make users willing to use such products again, which contributes to machine utilization improvement and medical system working efficiency. Focusing on the system thinking and process of how to identify the peak and end experiences in positive and negative aspects, this design practice also provides a valuable reference for enterprises to develop similar intelligent products.

2 User Experience in Smart Toilet

In recent 20 years, smart toilet industry has undergone tremendous innovation, both in technical aspect and in quality aspect. In this section, interactions between human and toilet are analyzed from the two dimensions of elements and process.

Many types of smart toilet brands and products are sold in the Chinese market, and the design elements related to interactions differ a lot. In Fig. 1, intelligent elements are summarized in five senses of sight, hearing, touch, smell, and perception. It is clear that the intelligent elements are diverse and complex, which has advantages and disadvantages. They bring plentiful functions to the users, but the burden in using process is correspondingly imposed.

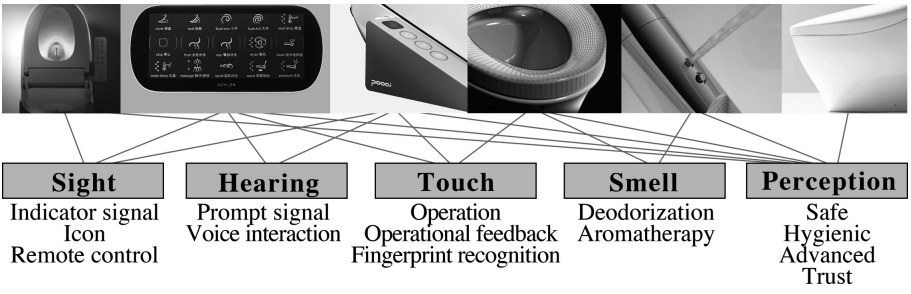


Fig. 1. Intelligent elements of smart toilet in different senses

In a default using process, users advance the process as the lead factor from human perspective, while data is calculated, analyzed, and presented accordingly from computer perspective and interface perspective. The flow of information between these three perspectives forms a complete closed loop: human provides data for computer, the computer provides analysis capability for interface, and the interface provides the visual information to the users and further influences user's behavior.

3 Identify Key Interactions Based on Peak-End Rule

3.1 Peak-End Rule

Peak-end rule, first proposed by psychologist Daniel Kahnman, means there are two main parts that can be remembered after a personal experience, that are exactly the feelings at the peak moment (negative and positive) and the end moment. It is shown in Fig. 2. This rule is based on the characteristics of the subconscious experience: only the experience at the peak and the end moments matters, and the proportion and duration of good or bad experiences have almost no effect on memory.

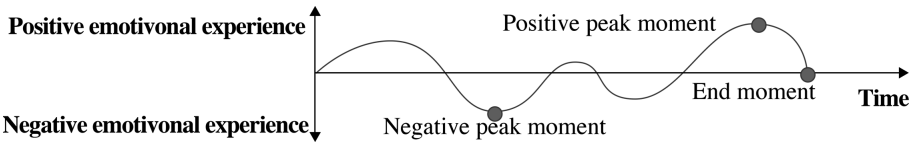


Fig. 2. Conceptual model of peak-end rule [11]

After identifying the peak and end moments, the corresponding user requirements can be improved to enhance user experience and satisfaction.

3.2 Identify Peak and End Interactions

To identify the peak and end moments, extensive and comprehensive interviews on 39 users who had experience in using smart toilets such as Panasonic, TOTO, Kohler, and Xiaomi were conducted to obtain the data needed to perform the study. Since Shanghai, Beijing, Hangzhou and Tianjin are the four cities with the highest penetration rate of smart toilets in China, users whose ages range from 20 to 60 in these four cities are the preferred research objects.

User Interview in the First Stage

In the first stage, we had about an hour communication with each participant, face to face or online to understand their age, gender, purpose, experience, expectations and specific operation steps.

As a result, the following results were found: (1) They hold an optimistic attitude about this smart product and expect to understand their even their family's health status through daily medical products. (2) The positive peak experience is missing: Most of the interviewed participants believe that the product does not give a particularly pleasant experience. On the contrary, many suggestions and expectations for improvement are proposed towards the confusing interactions. (3) The negative peak experience is appearing: The control panel is difficult to operate for the middle-aged and the elderly and most users are skeptical about the icons on the operator panel, in specific operations, for example, function startup and shutdown. This reflects a major problem. (4) Flushing is viewed as the last step and end experience.

Mainstream users are targeted at the same time the negative peak experience and end experience are initially identified. Based on the information of user interview, Table 1 lists nine functional requirements mentioned most by participants, which is to prepare for the next phase to identify the product's peak experience with Kano model.

Table 1. Sample of functional requirements

Factor	Functional requirements	Frequency
1	Get prompt for successfully testing	31 people
2	Sit comfortably and warmly	28 people
3	Trust the technology	26 people
4	Strengthen the sense of intimacy	20 people
5	Get visualized information in app quickly	20 people
6	Identify the icons on control panel	16 people
7	Choose different user mode	13 people
8	Use without instructions	10 people
9	Enjoy the stylish appearance	7 people

User Interview in the Second Stage

In the second stage, user interviews were conducted again with two purposes: To further verify the negative peak experience and end experience; To find functional requirements with expected quality and attractive quality combined Kano questionnaire. For example, ask participants "If the product offers the function, how do you feel? If not, how do you feel?" Give a value in −1 to 1 where "−1 means very dissatisfied, −0.5 means dissatisfied, 0 means indifferent, 0.5 means satisfied, 1 means very satisfied". Then you can know which quality this function belongs to.

As a result, the following results were found: (1) 38.4% and 23% of users respectively pointed out that the worst experiences are identifying the icons on control panel and sitting comfortably and warmly. It is similar to the results mentioned above. Therefore, the negative peak experiences are locked on these two key interactions. (2) Kano model proves that user satisfaction will only be enhanced when the product provides functions with expected quality or attractive quality, which is an important opportunity to find out positive peak experience. After the Kano questionnaire is conducted, we list five functional requirements included two expected requirements and three attractive requirements, which are showed in Table 2.

Table 2. Functional requirements with expected and attractive quality

Factor	Functions	Quality	SI	DSI
A	Choose different user mode	Attractive	0.42	−0.19
B	Sit comfortably and warmly	Expected	0.81	−0.73
C	Identify the icons on control panel	Attractive	0.70	−0.40
D	Get visualized information in app quickly	Attractive	0.62	−0.26
E	Get prompt for successfully testing	Expected	0.72	−0.82

In order to analyze the impact of different functional requirements, these five functional requirements are listed in the sensitivity analysis matrix of Fig. 3, where SI (Satisfaction Index) is the abscissa and DSI (Dissatisfaction Index) is the ordinate. The functions farther away from the origin have the greater sensitivity, that is to say Function B and Function E are the two most essential requirements. Function C, outside the radius circle, is also the negative peak experience, indicating that if ameliorate function C, the product not only can solve the pain point of the user experience, but also directly improve user satisfaction. Conversely, the functions within the radius circle are not sensitive and will not be developed.

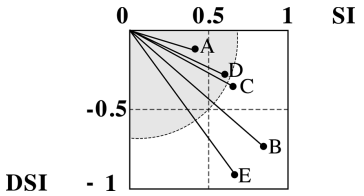


Fig. 3. Sensitivity analysis matrix for functions

Analysis in second phase shows that there are four key interactions based on peak-end rule. Peak experiences (positive and negative) include getting prompt for successfully testing, sitting comfortably and warmly and identifying the icons on control panel. End experience is flushing the toilet.

4 Improvement Design and Evaluation

For the four key interfaces identified with peak-end rule, we took the smart toilet produced by Beijing Geometry Technology Co., Ltd. as the specific research object, carrying out research and completing the design improvement. The following four items are done: (1) When data is successfully detected, the product emits apt tone. (2) Make toilet seat comfortable and constantly warm considering different groups of people. (3) There is text information on the control panel apart from the icons. (4) Upgrade manual flushing to automatic flushing.

Then 12 participants were invited to experience the improved products in Beijing and we recorded the completion time they took to finish a health check. Finally, they were asked to talk about their satisfaction with the toilet. Table 3 compares the duration and user satisfaction of smart toilet in two states.

Table 3. Comparison between original and improved products

Product	Average completion time of key task	Satisfaction
Original	33.92 s	33.3%
Improved	30.12 s	58.3%

Table 3 shows that the completion time of the improved product is significantly shorter than that of the original one, and the user satisfaction is also greatly enhanced.

5 Conclusion

The approach based on peak-end rule is proved to be efficient through user evaluation. Meanwhile, the application of peak-end rule in improving user experience of interactions between human and smart toilet can provide valuable reference for other intelligent products. There are many design factors that affect the user satisfaction and only four key interactions are identified and ameliorated in this paper, which improved user experience greatly. More design elements and functional requirements may be considered in the future.

References

1. Yang, M., He, H., Yuan, H.: Interaction design of products for the elderly in smart home under the mode of medical care and pension. In: 2nd International Conference on Human Aspects of IT for the Aged Population (ITAP) Held as Part of 18th International Conference on Human-Computer Interaction (HCI International). Lecture Notes in Computer Science, pp. 145–156 (2016)
2. Lallemand, C., Gronier, G., Koenig, V.: User experience: a concept without consensus? Exploring practitioners' perspectives through an international survey. Comput. Hum. Behav. **43**, 35–48 (2015)
3. Hassenzahl, M., Diefenbach, S., Goeritz, A.: Needs, affect, and interactive products - facets of user experience. Interact. Comput. SI, **22**(5), 353–362 (2010)
4. Baxter, K., Courage, C., Caine, K.: Understanding Your Users: A Practical Guide to User Research Methods, 2edn. Morgan Kaufmann, Burlington (2015)
5. Kujala, S., Roto, V., Vaananen-Vainio-Mattila, K.: UX curve: a method for evaluating long-term user experience. Interact. Comput. **23**(5), 473–483 (2011)
6. Xu, Y., Song, D., Gu, J.: Research on multi-fit design of toilet based on QFD and FBS model. Packag. Eng. **39**(24), 283–287 (2018)
7. Yoon, W.J., Shakir, M., Ali, Y.S.: Design and development of a smart multifunction toilet wheelchair (SMTW). In: 15th International Conference on Ubiquitous Robots (UR), Honolulu, HI, 26–30 June 2018, pp. 702–707 (2018)
8. Chung, H.A., Park, J., Hwang, J.B.: Design of ergonomic front-entry sitting toilet system for people using wheelchairs. J. Ergon. Soc. Korea **35**(5), 425–437 (2016)
9. Jin, D., Chen, H.: Research on home medical intelligent toilet based on ergonomics. Shandong Industrial Technology, vol. 128, no. 01 (2017)
10. Do, A.M., Rupert, A.V., George, W.: Evaluations of pleasurable experiences: the peak-and-end rule. Psychon. Bull. Rev. **15**(1), 96–98 (2008)
11. Fredrickson, B.L.: Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. Cogn. Emot. **14**(4), 577–606 (2000)

Research on Literature Searching Website Design Based on Usability Test

Haiwei Yan, Ruolin Gao, Yuanbo Sun(✉), and Bowen Sun(✉)

Beijing Institute of Technology,
5 South Zhongguancun Street, Haidian District, Beijing, China
hoiyanl992@163.com, l8810782763@163.com,
yuanbo@bit.edu.cn, doersun@gmail.com

Abstract. Based on the basic theory of usability test, a problem about how to improve the usability and UX (user experience) of literature searching websites are explored in this paper. Taking the CNKI website as an example, the usability test of the existing version of CNKI was carried out. The factors which affect the usability and UX of CNKI were obtained, and an iterative design were designed. Then, by comparing data of usability test between 2 versions. It is concluded that the optimized UX design can improve the usability of the website. Finally, key factors and guidance of optimizing the UX design of literature searching websites were proposed in this paper.

Keywords: Usability test · Website design · UX · t-test · Eye-tracking

1 Introduction

Nowadays, researchers can search the literature through literature searching websites easilier. There are many websites providing searching services such as Google scholar, Bing Scholar, CNKI (China National Knowledge Infrastructure), Elsevier Science-Direct, etc. Taking CNKI as an example, there are more than 40 million users of CNKI at present [1]. A questionnaire survey about the use satisfaction of the websites was conducted in this research. The results show that 83.7% of 600 users don't know how to conduct detailed search at the beginning of using; 93.8% of them have different problems using the websites. Therefore, there are defects of the existing literature searching websites' interface which affects the UX. Sandhu and Corbitt [2] proposed that the optimization of information acquisition process can promote the positive UX; Li [3] believes that UX designers need to focus on the integration of users' psychological and emotional research. Bang and Shaung [4] formed a website evaluation index based on UX and cognitive system. However, there are still some shortcomings. 1. Few research focus on the literature searching website layout design. 2. There are few usability analysis and iterative design for the website. In this research, the experiment is divided into 2 parts. The first part is used to explore the usability of the existing version of CNKI. The second part is used to explore the usability of the iterative design and compare it with the existing version.

These authors contributed equally to this study.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 104–108, 2020.
https://doi.org/10.1007/978-981-32-9244-4_13

2 Usability Test of the CNKI Existing Version

2.1 Literature Searching Website Interface

According to the survey, the home page layout of the website is divided into the following forms. The first form can be described as 'content-oriented' (No. 2 and No. 6 in Fig. 1) and the second form can be described as 'index-oriented' (No. 1, 3, 4, and 5 in Fig. 1). CNKI can be classified into 'content-oriented' form.

Fig. 1. Common document retrieval website layout form

2.2 Subjects and Experiment Content

In experiment of the existing version of CNKI, there were 10 subjects in the test, including 5 males and 5 females, all of whom were no visual abnormalities, aged 20–29 years old and use these websites 2–5 times per month. The test is divided into 3 parts: 1. Conducting user interviews to collect user information. 2. Asking subjects to complete the 8 tasks while tracking the eyes' movement, and record the task time of task1–task8 with stopwatch. 3. Filling out the SUS [5].

2.3 Web Page Lostness and Eye Movement Data Analysis

Fig. 2. Gaze plot - heat map - areas of interest

In Task 1, the eye movement trajectory is scattered (Fig. 2). Many other information distracts subjects. In Task 3, 5, 7, most subjects feel lost in pages. According to Smith's [6] evaluation formula for web page lostness, the rate of the lostness are calculated:

$$L = \text{sqrt}[\ (N/S - 1)2\ +\ (R/N - 1)2\ ]. \quad (1)$$

L: degree of "Lostness"; N: the number of different pages accessed (Unique Page-views); S: the total number of pages accessed (Pageviews); R: the minimum number of pages necessary to complete the task. The results are as follows: L3 = 0.87(>0.5); L5 = 0.68(>0.5); L7 = 0.6(>0.5). It indicates the subjects feel lost. The SUS score is 45, lower than the average, indicating that the website's usability is low.

3 Usability Test of the CNKI New Version

3.1 Subjects and Experiment Content

Some factors that are not conducive to UX are summarized and revised through experiment 1. They are embodied in the prototype and tested in experiment 2. Another 10 subjects who meet the requirements above are recruited in experiment 2 (Fig. 3).

Fig. 3. New version of CNKI prototype design.

3.2 Data Analysis

A Mann-Whitney U test was run to determine if there were differences in the task time of the 8 tasks (T1–T8) between the new version (GROUP 2) and the existing version (GROUP 1). Distributions of the task time for the new version and the existing version were not similar, as assessed by visual inspection. Task time for the new version in task 1, 2, 3, 5, 6, 7 were statistically significantly shorter than that for the existing version, which means new version can provide more efficient UX.

The analysis of data collected through the 2 experiments above are as follows (Tables 1 and 2):

Table 1. Mann-Whitney test-ranks of task time about the 2 versions of CNKI

	GROUP	N	Mean rank	Sum of ranks		GROUP	N	Mean rank	Sum of ranks
T1	1	7	13.14	92.00	T5	1	6	11.67	70.00
	2	10	6.10	61.00		2	10	6.60	66.00
	Total	17				Total	16		
T2	1	10	15.10	151.00	T6	1	9	14.00	126.00
	2	10	5.90	59.00		2	10	6.40	64.00
	Total	20				Total	19		
T3	1	7	14.00	98.00	T7	1	6	13.17	79.00
	2	10	5.50	55.00		2	10	5.70	57.00
	Total	17				Total	16		
T4	1	10	11.20	112.00	T8	1	10	11.90	119.00
	2	10	9.80	98.00		2	10	9.10	91.00
	Total	20				Total	20		

Table 2. Test statistics

	T1	T2	T3	T4	T5	T6	T7	T8
Mann-Whitney U	6.000	4.000	.000	43.000	11.000	9.000	2.000	36.000
Wilcoxon W	61.000	59.000	55.000	98.000	66.000	64.000	57.000	91.000
Z	−2.830	−3.477	−3.416	−.529	−2.061	−2.939	−3.037	−1.058
Asymp. Sig. (2-tailed)	.005	.001	.001	.597	.039	.003	.002	.290
Exact Sig. [2 * (1-tailed Sig.)]	.003b	.000b	.000b	.631b	.042b	.002b	.001b	.315b

a. Grouping Variable: GROUP. b. Not corrected for ties.

A SUS was used in the 2 experiments to evaluate their usability. 10 subjects completed experiment 1 (task A), and the other 10 subjects completed experiment 2 (task B). After completing the experiments, the 20 subjects scorded their SUS. The new version's SUS scores are compared with the existing version's SUS scores (Table 3). The two-sample t-test of SPSS is used for analysis (Tables 4 and 5).

Table 3. Group statistics of SUS scores

	GROUP	N	Mean	Std. deviation	Std. error mean
Score	1	10	45.000	13.1762	4.1667
	2	10	73.000	22.4475	7.0985

Table 4. Independent samples test of SUS scores

		Levene's test for equality of variances		t-test for equality of means			
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference
Score	Equal variances assumed	2.752	.114	−3.402	18	.003	−28.0000
	Equal variances not assumed			−3.402	14.544	.004	−28.0000

Table 5. Independent samples test of SUS scores

		t-test for equality of means		
		Std. error difference	95% confidence interval of the difference	
			Lower	Upper
Score	Equal variances assumed	8.2310	−45.2928	−10.7072
	Equal variances not assumed	8.2310	−45.5921	−10.4079

In this study an independent sample t-test were used to determine the user's discrepancies in the usability ratings of the old and the new versions of the CNKI interface. The study data did not have significant outliers, and it was close to a normal distribution within each group, and the variance was homogeneous. The results showed that the user scored the existing version (45.00 ± 13.18) lower than the new version (73.00 ± 22.45) with a difference of −28 (95% confidence interval −45.29–10.71). The results of the independent sample t-test indicate that t = − 3.402, P = 0.003, which indicates that the users has a statistical difference in the usability scores of the old and new versions. The usability of the new and existing versions is different.

4 Conclusions and Prospects

The new version (the index-oriented form) is significantly better in usability than the old version interface (content-oriented form). It can be concluded that the layout of the home page of the literature searching website should emphasize the main function "search". Reducing the image and operating steps can help improve the user's attention to the main function and reduce the lostness; this study can provide guidance for the website design. There are also some shortcomings in this paper: the sample size is relatively small, and more detailed functional tests are needed. In the further study, more detailed study of the website such as the study of the layout relationship between the website content and the main functions will be conducted.

References

1. Ming, D., Yuning, H.: Comprehensive evaluation of KDN knowledge discovery network platform based on user experience in CNKI. J. Inf. **32**(10), 182–187 (2013)
2. Sandhu, K., Corbitt, B.: User learning experience in web-based system: a case study, pp. 253–263. Springer, Heidelberg (2003)
3. Li, X.: Design of web user experience model based on universal psychological hierarchy theory. Inf. Doc. Work **1**, 62–66 (2010)
4. Bing, L., Shuang, L.: Research on comprehensive evaluation system of information quality based on user experience. Libr. Inf. Serv. **55**(22), 54–60 (2011)
5. Brooke, J.: SUS: a quick and dirty usability scale. In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, I.L. (eds.) Usability Evaluation in Industry. Taylor & Francis, London (1996)
6. Smith, P.A.: Towards a practical measure of hypertext usability. Interact. Comput. **8**(4), 365–381 (1996)

Lifetime Optimization of WSN Networks with Full-Coverage Nodes

Cong Chen(✉) and Jia Chen

Wuzhou University, Wuzhou, Guangxi, China
gxuwzchencong@qq.com

Abstract. Wireless sensor networks have been widely used for monitoring and tracking in industrial, rehabilitation medical, and environmental applications. In the network, sensor nodes are organized themselves into clusters with each one node having a fixed transmission range. These groups are called Sensor Covers (SCs), and the targets are monitored by the SCs. Typical target coverage algorithms often assume that the environment is known and each target is covered by only one node. However, these algorithms are not scalable. In fact, a target may require multiple node coverage, which is the K-coverage problem. For this reason, a K-Coverage Model Based Genetic Algorithm (KMGA) is proposed in this paper to increase the network lifetime. While satisfying the K-Coverage requirement, the KMGA model generates as many Sensor Covers (SCs) as possible, and then manages the Covers to ensure the switch between different Covers. Therefore, the network lifetime extends. Compared with similar algorithms, the performance of the proposed KMGA model in terms of network consumption and network lifetime is effectively improved.

Keywords: Sensor Covers · Lifetime optimization · K-coverage · KMGA model

1 Introduction

Typical target coverage algorithms assume that the environmental parameters are known, and these coverage algorithms find the optimal SCs to ensure that each sensor node covers a target [5, 6] to extend the network lifetime. Yet if the sensor nodes fail, such as running out of energy, these algorithms are no longer scalable. Therefore, a target may need to be covered by more than one sensor node. With multiple sensor nodes covering the target at the same time, even if a node fails, it can still maintain the continuous coverage of the target.

Covering the target is a time-consuming and energy-consuming process. Besides energy, there are still some optimization constraints, such as time consumption, continuous coverage, and cover structure [7]. These problems are called K-coverage problems. This requires any target to be covered by at least K sensors node [8]. However, it is an NP-hard optimization problem, and this problem can be solved by metaheuristic optimization [9].

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 109–114, 2020.
https://doi.org/10.1007/978-981-32-9244-4_14

Literature [10] proposed the flow decomposition algorithm (FDA) and compared it with the fixed direction sensing arrangement problem. The purpose of the FDA algorithm is to decompose the maximum flow into multiple single flows, and each single flow represents the path from the source node to the sink. At the same time, all sensor nodes on this path form a cover.

This paper proposes a K-coverage algorithm based on the GA model to maximize the network lifetime and decides which SCs are active to complete the current data transfer round. In other words, GA is used to optimize the coverage requirements of WSNs and provide as continuous monitoring of the target area as possible. In brief, the research objectives of this paper are as follows: (1) Firstly, a GA-based coverage model is proposed to generate as many SCs as possible to meet the K-coverage requirements; and (2) Coverage management algorithms are used to switch the active states of different SCs so as to maximize network lifetime. Experimental data shows that the proposed KMGA model effectively extends the network lifetime.

2 KMGA Model

The network model KMGA framework is shown in Fig. 1. This model consists of three main phases. In the first stage, all sensor nodes in the network are coded using two-bit chromosomes. Then, in the second stage, the base station runs the GA algorithm, selects the optimal number of SCs, and then forms the Covers according to the sensing range and the target position of each sensing node. Finally, each stain goes to assess whether it guarantees that all targets are covered. Based on the possible number of acquired Covers, the expected energy consumption of each sensor node will be calculated. And then the base station determines which Covers remain active and thus maintain maximum network lifetime.

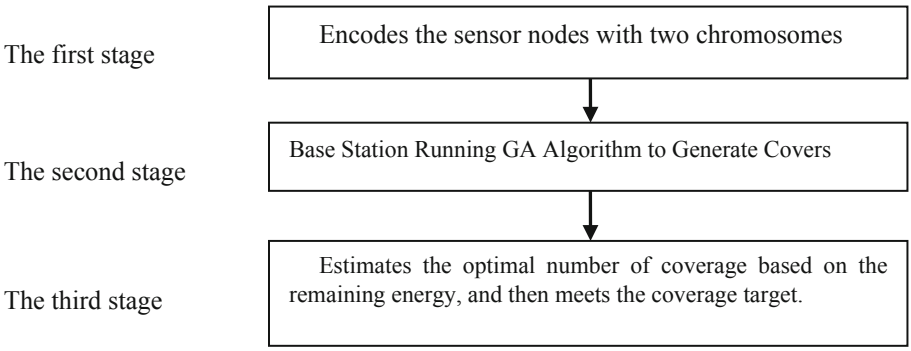


Fig. 1. KMGA model

3 GA Algorithm

In the KMGA model, each gene in each chromosome represents a sensor node. The value of this gene may be 1 or 0, where 1 indicates that the node is active and 0 indicates a dormant state.

The GA algorithm generates chromosomes through crossover and mutation operations and estimates the fitness function. In the KMGA model, the graceful function of GA consists of the remaining energy $\tilde{E}$, the overall expected energy consumption ΔE, and the distance between the sensor node and the sink $d(s, B)$.

Therefore, using Eq. (2), the energy cost by each sensor node is calculated and the expected energy consumption ΔE is estimated. Finally, the moderate function is shown in Eq. (4):

$$f = \frac{\tilde{E}}{NE(0)} + \frac{E'}{\Delta E} + \frac{1}{\sum_i d(s_i, B)} \tag{4}$$

In the formula (4), E' represents the total energy consumed when the sensing node transmits a message to the base station in a direct manner. N is the total number of sensor nodes.

4 Simulation Experiment and Analysis

This section established some experimental evaluations of the KMGA model applied to different K-coverage situations. In addition, the energy consumed by each node of the transmission wheel was analyzed. Table 1 lists the network parameters, among which E_{fs} represent free space energy consumption and E_{mp1} represent the energy needed for mobile sensing.

Table 1. Simulation parameters

Coverage	Method	100 m × 1 00 m		200 m × 200 m	
		FTU	LTU	FTU	LTU
K = 1	FDSSP	675	1240	509	914
	FDA	650	1301	630	1000
	VP-NL	751	1514	689	1225
	Proposed	1050	1847	886	1547
K = 2	FDSSP	520	839	414	620
	FDA	487	992	391	785
	VP-NL	586	1053	421	869
	Proposed	864	1375	627	1148
K = 3	FDSSP	312	641	209	300
	FDA	383	787	301	417
	VP-ML	398	803	210	524
	Proposed	632	1015	402	728

(1) This experiment considers three different K-coverage situations with K equal to 1, 2 and 3. It measures the time duration between when the first target becomes uncovered (FAT) and when the last target becomes Uncovered (LTU). Experimental data is shown in Table 2.

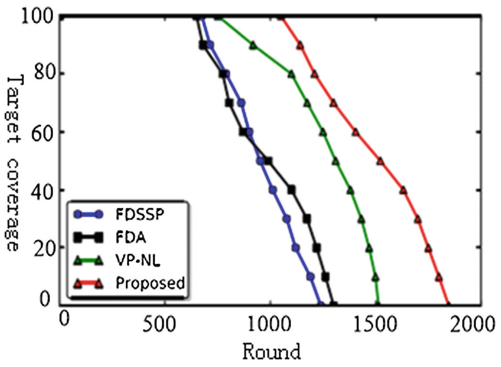
Table 2. Network life(s)

Properties		Values
Number of nodes		100
Initial node energy		0.5 J
Idle state energy		50 nJ/round
Data aggregation energy		5 nJ/bit
Amplification energy	$d \geq d_0$	10 pJ/bit/m^2
(cluster head to base-station)	$d < d_0$	0.0013 pJ/bit/m^2
Amplification energy	$d \geq d_1$	$E_{fs}/10 = E_{fs1}$
(node to cluster head)	$d < d_1$	$E_{mp}/10 = E_{mp1}$
Packet size		400 bits

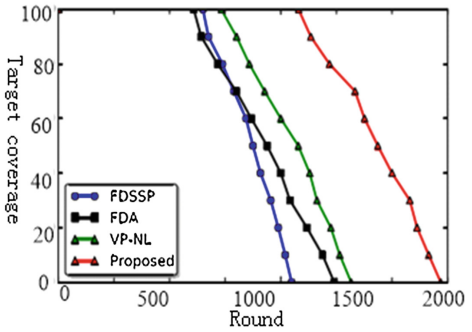
From Table 2, it can be seen that as the coverage level increased (K value increased), energy consumption was accelerated, resulting in a decrease in FTU and LTU. In other words, the network life was inversely proportional to the K value. In addition, compared with FDSSP [15], FDA [10], and VP-NL [5], the network life of KMGA was improved. The reason is that the KMGA model balances energy consumption and thus ensures that all sensor nodes consume energy at the same level.

(2) This experiment considered two cases: (1) 100 nodes are distributed at 100 m $\times$ 100 m, covering 10 targets; (2) 200 nodes are distributed at 200 m $\times$ 200 m, covering 20 targets. The experimental data for these two cases were shown in Figs. 3 and 4, respectively.

From Fig. 2, it can be seen that the target coverage of the proposed KMGA algorithm was higher than that of FDSSP, FDA, and VP-NL in the three cases with K equal to 1, 2. For example, when K = 1 at that time, when the number of rounds went through to 1500, the coverage of FDSSP, FDA, and VP-NL was zero, while the target coverage of the KMGA algorithm still reached 60%.



(a) K=1



(b) K=2

Fig. 2. Target coverage (Case 1)

5 Conclusion

Aiming at the target supervision application of wireless sensor network, a K-coverage model KMGA based on genetic algorithm was proposed. With the KMGA model, it is ensured that each target was covered by multiple sensor nodes at the same time. Multiple Covers were generated as much as possible through the GA algorithm. Then, while guaranteeing coverage requirements, the coverage between Covers was switched based on the remaining energy of the nodes, and it resulted in thereby extending the network lifetime. The experimental data shows that the proposed KMGA model can effectively extend the network lifetime and ensure coverage.

Project Funding. Guangxi Colleges and Universities Key Laboratory of Professional Software Technology (Wuzhou University);

Promotion of basic ability of young and middle-aged teachers in Universities in Guangxi (2018KY0539);

References

1. Alvi, A.N., Bouk, S., Ahmed, S.H., Javaid, N., Kim, D.: Enhanced TDMA based MAC protocol for adaptive data control in wireless sensor networks. Int. J. Commun. Netw. **17**(3), 247–255 (2015)
2. Min, F., Sijia, X.: Location routing improved algorithm based on cavity model. J. Sens. Technol. **25**(11), 1556–1561 (2012)
3. Stine, J.A.: Exploiting smart antennas in wireless mesh networks using contention access. IEEE Wireless Commun. **13**(2), 38–49 (2016)
4. Ferentinos, K.P., Katsoulas, N., Tzounis, A., Bartzanas, T., Kittas, C.: Wireless sensor networks for greenhouse climate and plant condition assessment. Biosyst. Eng. **53**(6), 70–81 (2017)
5. Yang, Q.: Variable-power scheduling for perpetual target coverage in energy harvesting wireless sensor networks. In: Proceedings of International Symposium on Wireless Communication Systems, pp. 281–285 (2015)
6. Mohd Alia, O.: Dynamic relocation of mobile base station in wireless sensor networks using a cluster-based harmony search algorithm. Inf. Sci. **45**(385), 76–95 (2017)
7. Ahmed, S.H., Kim, D., Bouk, S., Javaid, N.: Error control based energy minimization for cooperative communication in WAN. ACM SISAPP Appl. Comput. Rev. **14**(3), 55–64 (2014)
8. Yang, Q., He, S., Li, J., Chen, J., Sun, Y.: Energy-efficient probabilistic area coverage in wireless sensor networks. IEEE Trans. Veh. Technol. **64**(1), 367–377 (2015)
9. Elhoseny, M., Farouk, A., Zhou, N., Wang, M.-M., Abdalla, S., Batle, J.: Dynamic multi-hop clustering in a wireless sensor network: performance improvement. Wireless Pers. Commun. **4**(8), 1–21 (2017)
10. Wan, X., Wu, J., Shen, X.: Maximal lifetime scheduling for roadside sensor networks with survivability. IEEE Trans. Veh. Technol. **64**(11), 5300–5313 (2015)
11. Metawa, N., Kabir Hassan, M., Elhoseny, M.: Genetic algorithm based model for optimizing bank lending decisions. Expert Syst. Appl. **3**(80), 75–82 (2017)
12. Elhoseny, M., Yuan, X., Yu, Z., Mao, C., El-Minir, H., Mohamed Riad, A.: Balancing energy consumption in heterogeneous wireless sensor networks using genetic algorithm. IEEE Commun. Lett. **19**(12), 2194–2197 (2015)
13. Ebrahimian, N., Sheramin, G.Y., Navin, A.H., Foruzandeh, Z.: A novel approach for efficient k-coverage in wireless sensor networks by using genetic algorithm. In: Proceedings of International Conference on Computational Intelligence and Communication Networks, pp. 372–376 (2010)
14. Mnasri, S., Thaljaoui, A., Nasri, N., Val, T.: A genetic algorithm-based approach to optimize the coverage and the localization in the wireless audio-sensors networks. In: Proceedings of International Symposium on Networks, Computers and Communications, pp. 1–6 (2015)

Proceedings Design and Implementation of Cultural Communication and Promotion System Based on Big Data

Xuan Qi[1(✉)] and Bo Ding[2]

[1] China-Korea Institute of New Media,
Zhongnan University of Economics and Law, Wuhan, Hubei, China
shixuan77@163.com
[2] China Everbright Bank Wuhan Branch, Wuhan, Hubei, China

Abstract. In the era of big data, the status quo of continuous change of information has brought about great changes in the mode of public acceptance, cognition and dissemination of information. In the mass change to accept information, cultural space construction should be accordingly strengthened new ways of transmission channel, this paper designed and implemented a kind of new culture promotion system, and put forward the recommendation method based on the culture content, it according to the user the choice or the like of the object (product), recommend similar options for the user. Content-based recommendation does not need to be based on the user's evaluation of the project, but more needs to use machine learning method to get the user's interest information from the case of the description of the content characteristics, to find the right efficient media, the right communication method.

Keywords: Cultural communication · Data mining · Personalized recommendation · JSP

1 Introduction

Cultural communication and extension from archaeology, development, research and design to the final face of society [1, 2]. Forming humanistic output is a complete information chain transmission link, and the absence of any link will indeed affect the final effect of cultural communication. Cultural display space from the content layout to display design are affecting the efficiency of display and communication, through various channels of publicity impact on the breadth and depth of the final cultural promotion. The cultural communication effect brought by the visitor flow is also an important way to promote the exhibition, which is complementary to each other.

Personalized recommendation system is a kind of advanced business intelligence platform based on mass data mining to help various websites provide users with fully personalized decision support and information services. For the cultural communication and promotion system, It can helps to collect user characteristics information and make personalized recommendations for users according to user characteristics, so as to meet the personalized needs of customers [3, 4]. The cultural communication and promotion

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 115–120, 2020.
https://doi.org/10.1007/978-981-32-9244-4_15

system developed in this paper adopts the JAVA development environment, USES the MVC development mode, takes MySQL as the system database and Tomcat as the server. The whole system is easy to operate, user-friendly, flexible and practical [3].

2 Analysis and Design of Cultural Communication and Promotion System

Modern forms of cultural communication are presented in a highly concentrated state. Massive and multi-class information hierarchy needs to be expressed through limited transmission channels. As a carrier, any kind of cultural communication and promotion system will face bearing capacity considerations.

2.1 Demand Analysis of the System

System demand analysis is an indispensable link of platform development, in order to make the system better and more perfect to be designed, it must be investigated first. On the basis of system investigation, the function of the new system is analyzed in detail, so as to develop a complete system design. Through past browsing and reference to some cultural transmission system, as a cultural transmission system, should be divided into front desk module and the background management module, front desk module should include login, registration, shopping cart, order, such as module, the background module should be included, the administrator login, add classification, add goods, check the order module. According to the requirements of the system step by step, planned development. On the other hand, the developed system should meet the following standards:

(1) The system interface is user-friendly and easy to operate.
(2) The system should be easy to maintain and expand system functions.
(3) To ensure that the system can accommodate a certain number of customers to access at the same time, to ensure the stability and robustness of the system.
(4) The relationship between background databases is clear to avoid data redundancy.

2.2 Architecture Based on B/S

Any form of carrier, will face bearing capacity considerations. From the era of “we media” to the current era of big data, receptors are actively or passively receiving all kinds of information, and publishers have the obligation to optimize the process of receiving and digesting information.

In the B/S architecture system, users send requests to many servers distributed on the network through the browser. The server processes the requests of the browser and returns the required information to the browser [4]. The B/S structure simplifies the work of the client, which requires only a small amount of client software to be configured. The server takes on more work, with database access and application execution done on the server. The browser makes the request, and the rest of the work, such as

data requests, processing, result returns, and dynamic page generation, is all done by the Web Server [5].

In fact, the B/S architecture is to separate the two-layer C/S structure transaction logic module from the client's task, and the Web server forms a separate layer to bear its task, so that the client's pressure is reduced and the load is distributed to the Web server. This three-tier architecture is shown in Fig. 1.

Fig. 1. Diagram of three-tier B/S architecture

3 Research on Recommendation Algorithm Based on Collaborative Filtering

There are many objective defects and disadvantages in the traditional cultural communication and promotion systems. Due to the monotonous and passive form of the traditional "listing" display, visitors can only receive knowledge in a monotonous and passive way. There is no purpose in receiving information, and the effect of receiving information cannot be fed back. However, information recommendation and modern digital interactive technologies have changed the traditional deadlock.

Content-based recommendation is the earliest recommendation method, which recommends similar choices for users based on the objects (products) they have chosen or liked in the past. Content-based recommendations do not need to be based on the user's comments on the project, but more need to use machine learning method to get the user's interest information from the case of content feature description. The content-based recommendation process generally has three steps, as shown in Fig. 2:

(1) Describe the object, extract object features to represent the object;
(2) Feature learning: the user's preferences can be learned from the selected object feature data in the past;
(3) Make recommendations, compare the characteristics of users' preferences with those of candidate objects, and recommend a group of objects with the greatest relevance to users.

The disadvantages of content-based recommendation methods are:

(1) Feature extraction is difficult;
(2) Unable to dig out the potential interests of users;
(3) Unable to generate recommendations for new user.

User-based collaborative filtering is by far the most successful personalized recommendation technology in practical application. The basic idea of the algorithm is to recommend items of interest to target users who share the same hobbies, as shown in Fig. 3.

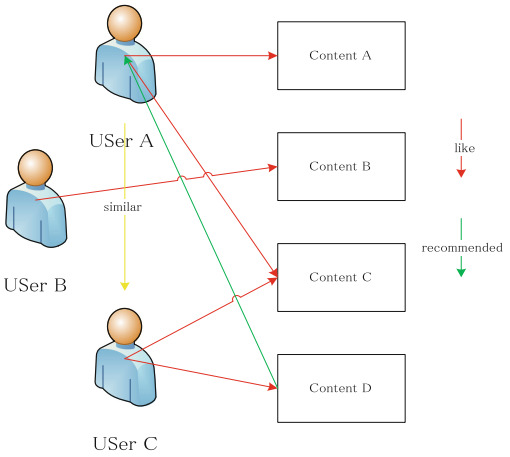


Fig. 2. The basic process of CB

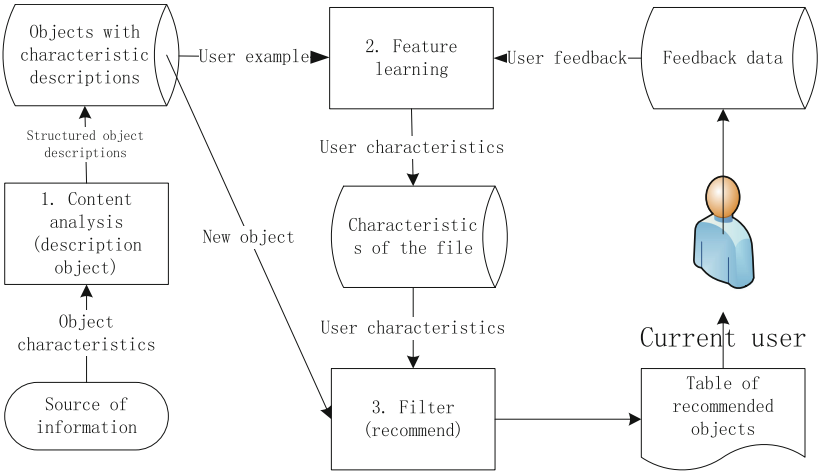


Fig. 3. The basic idea of collaborative filtering based on users

The main tasks of user-based collaborative filtering recommendation algorithm are user similarity measurement, nearest neighbor query and prediction score. This paper adopts the method of similarity between users.

Cosine similarity: user-item scoring matrix can be seen as a vector in the n-dimensional space, for the items without scoring will be set to 0, cosine similarity measurement method is to measure the similarity between users by calculating the cosine Angle between vectors. Let the vectors I and j represent the scores of user I and user j in n-dimensional space respectively, then the similarity between user I and user j

is studied using the e-commerce personalized recommendation algorithm based on collaborative filtering as follows:

$$\text{sim(i, j)} = \frac{\sum_{c\in I_{i,j}} \left(R_{i,c} - \bar{R}_i\right)\left(R_{j,c} - \bar{R}_j\right)}{\sqrt{\sum_{c\in I_i} \left(R_{i,c} - \bar{R}_i\right)^2}\sqrt{\sum_{c\in I_j} \left(R_{j,c} - \bar{R}_j\right)^2}} \tag{1}$$

After getting the nearest neighbor of the target user, the corresponding recommendation results will be generated. The calculation formula of user u's prediction score for item I is as follows:

$$P_{u,i} = \bar{R}_u + \frac{\sum_{n\in NN_u} sim(u,n) * \left(R_{n,i} - \bar{R}_n\right)}{\sum_{n\in NN_u} |sim(u,n)|} \tag{2}$$

The most typical measures of accuracy are MAE (mean absolute error), RMSE (mean square error), and NMAE (standard mean error). Their calculation forms are shown below.

$$\text{MAE} = \frac{1}{n}\sum_{a=1}^{n} |p_{ia} - r_{ia}| \tag{3}$$

$$\text{RMSE} = \sqrt{\frac{1}{n_i}\sum |p_{ia} - r_{ia}|^2} \tag{4}$$

$$\text{NMAE} = \frac{MAE}{r_{max} - r_{min}} \tag{5}$$

Where n is the number of user I rating products in the system, p_{ia} and r_{ia} are respectively the predicted rating and the actual rating. N_i is the number of user-product pairs in the system. r_{min} and r_{max} are the minimum and maximum values of the user rating interval, respectively.

4 Conclusion

Our country in recent years, the construction process of all kinds of cultural exhibition space to promote the rapid, technology innovation is instant, but continues to progress at the same time, inevitably, some common problems, with the era of big data, the present situation of the information change constantly, the public accept and cognitive mode and dissemination of information, there is a huge change. In the mass change to accept information, cultural space construction should be accordingly strengthened new ways of transmission channel, this paper designed and implemented a kind of new culture promotion system, and put forward the recommendation method based on the culture content, it according to the user the choice or the like of the object (product), recommend similar options for the user. Content-based recommendation does not need to be based on the user's evaluation of the project, but more needs to use machine

learning method to get the user's interest information from the case of the description of the content characteristics, to find the right efficient media, the right communication method.

References

1. Beaton, L., Freeman, R.: Oral health promotion and homelessness: a theory-based approach to understanding processes of implementation and adoption. Health Educ. J. **75**, 184–197 (2016)
2. O'Neil, M.E., Fragala-Pinkham, M., Ideishi, R.I., et al.: Community-based programs for children and youth: our experiences in design, implementation and evaluation. Phys. Occup. Ther. Pediatr. **32**, 111–119 (2012)
3. Downey, A.M., Frank, G.C., Webber, L.S., et al.: Implementation of "Heart Smart:" a cardiovascular school health promotion program. J. Sch. Health **57**, 98–104 (2010)
4. Levin-Zamir, D., Keret, S., Yaakovson, O., et al.: Refuah Shlema: a cross-cultural programme for promoting communication and health among Ethiopian immigrants in the primary health care setting in Israel: evidence and lessons learned from over a decade of implementation. Glob. Health Promot. **18**, 51–54 (2011)
5. Gaalen, J.L.V., Bakker, M.J., Bodegom-Vos, L.V., et al.: Implementation strategies of internet-based asthma self-management support in usual care: study protocol for the IMPASSE cluster randomized trial. Implementation Sci. **7**, 113 (2012)

Discussion on Digitization Construction of Cangyuan Rock Paintings in the New Media Era

Qian Xu(✉)

China Korea School of New Media,
Zhongnan University of Economics and Law, Wuhan 430072, China
243043454@qq.com

Abstract. Cangyuan Rock Paintings is one of the most ancient rock paintings found in ethnic minority areas in China so far, representing the most objective record of material and cultural life and spiritual life in a certain period of the Va nationality. There are numerous themes in Cangyuan Rock Paintings which implicate the essence of Chinese national culture and reflect the culture deposits and inheritances of Chinese people for thousands of years. In the era of digitization, network and electronic technology have penetrated into various aspects of our daily life. Therefore, it is imperative for us to apply digital technology to the protection and dissemination of cultural heritages. This paper places the Cangyuan Rock Paintings into the background of protection and dissemination of cultural heritages, taking advantage of digitization to find a feasible and practicable way for the dissemination and survival of Cangyuan Rock Paintings.

Keywords: New media · Cangyuan Rock Paintings · Digitization construction · Cultural heritage · Digital transmission

1 Summary of Cangyuan Rock Paintings

Rock paintings is the prophase of paintings, which vividly represents the cultural memory of some ethnic groups. With a history of more than three thousand years, Cangyuan Rock Paintings is one of the ancient rock paintings which is located in Cangyuan Va Autonomous County, Yunnan Province. Cangyuan Rock Paintings is the most concentrated one which have the maximum images among more than fifty rock paintings in Yunnan. Recording not only various scenes of people's life at that time but also the development process of the Va nationality, it is a significant part of brilliant Chinese culture as well as the spiritual wealth shared by the Wa people in China.

From the perspective of artistry and style, Cangyuan Rock Paintings is mainly based on the "flat-painted silhouette" style, highlighting the characteristics of the image without sticking to details. Characters mostly apply inverted triangle to represent trunk, circular to represent head, and lines to represent limbs. In the meantime, there are more males and society, the agricultural production scene was not shown.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 121–127, 2020.
https://doi.org/10.1007/978-981-32-9244-4_16

From the perspective of value and worth, Cangyuan Rock Paintings is the vivid reflection of fewer females in the paintings. Characters are mostly positive while animals are mainly lateral, which is consistent with the requirements of Egyptian art for the law of frontality of characters. The color of the paintings is mainly monochrome, particularly red, for the reason that hematite power or cow blood were evenly dyed on it with fingers and feathers taking advantage of the symbolic meaning of red color in order to express the mystery of rock paintings, which is as the same phenomenon that red is preferred in primitive art.

From the perspective of expression content, Cangyuan Rock Paintings is divided into hunting and gathering scenes. There are also dancing pictures, villages and buildings, which can be recognized as stilt style architecture. Meanwhile, there are scene pictures such as war, boating, sacrifice, etc. Ideographic symbols and decorative patterns are also widely used, such as villages, caves, the moon, the sun, handprints and so on. Since it has not yet entered the agricultural social life in this region and the encyclopedia of lives of local ancestors, which results in its significant historical materials value and cultural value. In the meantime, it provides visual specimen for the research of understanding the spiritual world and national cultural of ancient Chinese.

2 Significance of Digitization Construction of Cangyuan Rock Paintings

Cangyuan Rock Paintings has rich and distinctive cultural connotations. Enhancing the digital protection and dissemination of Cangyuan Rock Paintings play an important role in the dissemination of telling Chinese stories and Yunnan's history and culture. However, the current propagation mode is insufficient to express, excavate and disseminate the cultural connotation of Cangyuan Rock Paintings, while digital communication can better express it from the aspects of verisimilitude, interaction and interest.

2.1 Digitization of Cangyuan Rock Paintings Is the Irresistible Trend of Heritage Protection

With the rapid development of new technology and the success experience accumulated in "Digital Dunhuang" and "Digital Palace Museum", it is possible for the digital protect and presentation of Cangyuan Rock Paintings, which provides a new method for the research, renovation and visual simulation of the century heritage. The digital preservation, reproduction and dissemination of Cangyuan Rock paintings can be realized by adopting digital methods such as scanning, photography, digital editing, three-dimensional animation, virtual reality and network to collect data, process data and disseminate information. Compared with the uniqueness, non-sharing and non-renewability of concrete objects, digital data storage is infinite, shared and renewable, and can be permanently preserved and used. "Digital Cangyuan Rock Paintings" is the meaning of implementing the digitization of world cultural heritage. Through realizing the digitization of Cangyuan Rock Paintings we can not only obtain permanent archives for protection as well as provide detailed information for heritage research, but

also develop digital programs for tourism education and cultural dissemination, so as to better protect and disseminate Cangyuan Rock Paintings, and ultimately realize the harmonious integration of cultural value and commercial value. Applying new media technology in the combination of dissemination can promote the formation pf a new industrial chain of scientific research, culture, tourism and education, and achieve the unity of social and economic benefits.

Vividly present Cangyuan Rock Paintings to enhance the overall awareness of the audience.

Technology reproduction is an urgent problem that Cangyuan Rock Paintings has a large volume and wide pattern distribution. Through virtual reality technology, sound photoelectric effect and story interpretation, we can vividly show the complete style and features of Cangyuan Rock Paintings. For example, with the development of live high definition fulldome digital movies of which the theme is cultural heritage, and through the latest 8K technology together with musics, commentaries and subtitles, the audience in the virtual field can feel as if they were in the scene. It can not only enhance the psychological feelings of the audience but also make them feel shock from the bottom of heart which can not be realized by normal propagation medium and means such as real scene inspections and photo exhibitions. Being personally on the scene, people have a more profound understanding of Cangyuan Rock Paintings while being deeply impressed by the superlative craftsmanship of ancient people. Affected by the national wisdom and cultural connotations displayed by the sight, people's attitude of Cangyuan Rock Paintings change from unwilling to know to willing to positively understanding and disseminating it. Through the deep inner experience of digital technology, the cultural communication develop to a deeper level.

2.2 Digital Communication Has Breadth and Depth, Which Promotes Cangyuan Rock Paintings to "Go Out"

The communication revolution in the new media era makes information dissemination more diversified, three-dimensional, real-time and interactive. Virtual reality technology and the mode of mobile terminal communication are more in line with the media usage habits of young people. Making full use of digital technology to enrich the display form and content, we are able to explore the wisdom and customs of the ancients, as well as tell the story and historical legends of Cangyuan Rock Paintings. Based on the rich connotation and unsolved mysteries, we can develop various forms of digital animation games, singing stories, literary works, artworks and so on, so that Cangyuan Rock Paintings can be accepted by young audience through digital dissemination and also go abroad, ultimately go to the world. Promoting the overall understanding of audience of Cangyuan Rock Paintings is conductive to enhancing national pride and expanding the scope of dissemination and influence of excellent national culture in China.

3 The Way of Digitization Construction of Cangyuan Rock Paintings

In the new media era, technological innovation has brought new opportunities for the digital dissemination of Cangyuan Rock Paintings. Based on the opportunities, the digital dissemination should break the limitation of time and space on the basis of the digitization of Cangyuan Rock Paintings heritage in order to meet the multiple requirements of cultural dissemination, tour appreciation, leisure and entertainment, and academic research.

3.1 Apply Digital Technology to Build "Digital Cangyuan Rock Paintings"

3.1.1 Illuminant

The commonly used scanners currently are flatbed scanners and roller scanners. The former uses fluorescent tube as illuminant, and parts other than the manuscript are also irradiated. The refraction of these light will reduce the contrast of dark part and as a result, affect the level of its reproduction. The latter scanner applies point scanning while not being affected by excess light. Pointolite with high intensity will focus on the manuscript so that we can sweep more details in the dark part and improve image contrast. However, we would still use flatbed scanners instead of roller scanners for the latter need to paste the manuscript on itself which is unrealistic for painting and calligraphy works of art.

In order to insure the effect of scanning, we make use of LED standard illuminant. In the meantime, we can effectively avoid the influence of ambient light and background color on color and level by guaranteeing that the workplaces where the paintings and calligraphy are scanned are closed and the walls are gray or black.

3.1.2 White Balance

The color rendering of illuminant would change to some extent when using time changes. Thus, we need to set white balance regularly and preheat the illuminant for around thirty minutes before scanning.

3.1.3 Focalize

By manual or automatic focusing, the image can be placed on the image sensor to ensure that the scanned image is clear and not blurred.

3.1.4 Color Management

Making special standard color target to generate properties file aimed at the reproduction of artworks. Forming ICC color management feature files which is suitable for art reproduction to achieve real "see as you get".

3.1.5 Resolution Ratio

It is significant to select the appropriate resolution ratio for digital images are composed of countless pixels. Generally speaking, the legibility of the image increases with the scanning resolution while scanning resolution ratio is restricted by many factors and is

not that the higher the better. Excessive precision will prolong the scanning time, occupy too much memory, and slow image processing speed; or when the scanning resolution ratio is higher than the resolution ratio of the output device, additional pixels would be discarded in the printing process, which can also reduce the image legibility.

3.2 Skillfully Use Mobile Terminal to Expand Communication Channels

Up to December 2017, the scale of Chinese netizen has reached 772 million in which mobile netizen takes up a large slice of 753 million. Mobile phones continue to occupy the use of other personal Internet devices. Intelligent devices with mobile phones as the principle part has become the basis of "all things interconnect with each other". Served as a media platform, micro-blog and WeChat combine mass communication, group communication and interpersonal communication together. Intelligent devices make it quickly and instantly for "micro-platform" to disseminates information including text, picture, audio and video, etc., satisfying the information requirements from various viewpoints and levels as well as changing people's lifestyle and information dissemination mode, which brings about unprecedented experience of convenient communication.

Propaganda departments or traditional medium should push news and information about Cangyuan Rock Paintings based on "two micro-one end" and other "micro-platforms", timely and vividly disseminate Cangyuan Pock Paintings and its culture in various forms such as text, picture, audio and video. For example, the WeChat Subscription of Yunnan News covered a lot on Cangyuan Rock Paintings on its official WeChat Subscription, which guided the audience to make comments meanwhile promoted their awareness of Cangyuan Rock Paintings.

Creating mobile phone client (APP) about Cangyuan Rock Paintings, updating content continuously according to audience's habits and requirements and upgrading application software can not only strengthen habitual behavior and increase page view, but also consolidate the scale of audience group size, so that the audience have strong browsing desire, so as to more effectively realize "residence" and "solicitation". Currently, APPs that are related to Cangyuan Rock Paintings are in the minority, most of which are APPs of a certain medium who add reports and information about Cangyuan Rock Paintings to their own APPs, or open up an specialized area in APP to introduce Cangyuan Rock Paintings. Although this method can produce a certain communication effect for a period of time, it is difficult to form a lasting communication force, so we still need to create an own APP of Cangyuan Rock Paintings in order to help people deeply understand it. Through intelligent technology and "two micro-one end" and other "micro platforms", we can better enhance the audience's understanding of Cangyuan Rock Paintings as well as spread it more effectively.

3.3 Tell a Great New Story of Cangyuan Rock Paintings Through Animation, Game and Micro-film

To disseminate Cangyuan Rock Paintings and tell its stories well, we can entrust animation production companies to digitize Cangyuan Rock Paintings stories with rich connotations by means of digital animation art forms as well as present and disseminate

them to the general audience through digital means. Animations, as a new method for culture spreading, this cultural industry and its export are conducive to enhancing the comprehensive strength of Chinese culture. On the other side, as a medium, it has become a platform for the spread of traditional culture of our nation that can reflect the spirit of Chinese national art and culture in front of the world more vividly.

As the pace of life speeds up, the "fragmented" information receiving mode of micro-film adapts to the real-time consumer demand of the audience. Producing microfilms based on the background or by mixing elements of Cangyuan Rock Paintings, the mysterious ancestor culture of Yunnan Province is disseminated through popular stories, artistic means of expression, and well-structured narrative plots. The mysterious ancestor culture is familiar to the audience in an acceptable way, which opens up the public's understanding of Cangyuan Rock Paintings. For the micro-film is short and concise, it can convey the intended meaning rapidly in a limited time. By showing the culture of Cangyuan Rock Paintings through micro-film, the wisdom of the ancestors of Yunnan Province can be presented in the way of video recording. Through that way, we are able to disseminate the culture of Cangyuan Rock Paintings in storytelling and tell the stories of it.

In the meantime, as an important way of leisure for mobile phone users, games have gathered a large number of young audiences. By adding some elements of Cangyuan rock Paintings into games, taking the rich connotation and unsolved riddles of it as the material, and developing games suitable for teenagers' entertainment habits, we can not only improve the recognition of Cangyuan Rock Paintings among young people, but also effectively promote the audience to understand the excellent Cangyuan Rock Paintings.

Under the background of "going out" of Chinese culture, new forms of dissemination such as animation, micro-films and games can often "moisten things silently". To some extent, Cangyuan Rock Paintings have gained wide attention no matter in the industry or academia; however, out of Yunnan Province, nationwide even worldwide, we may only know that Cangyuan Rock Paintings is a precious cultural heritage while know nothing about the cultural connotation of it, which is the bottleneck of the current means of dissemination. Animation, micro-film and game are concerned by young audiences. Through adding Cangyuan Rock Paintings elements in them, the audiences are able to contact and accept Cangyuan Rock Paintings in a subtle way, and even be interested in it as well as participate in the dissemination. They are both receivers and disseminators, which is conducive to forming the audiences' deep and overall understanding of Cangyuan Rock Paintings.

4 Conclusion

Digitization provides a favorable platform for the protection of cultural heritage of Cangyuan Rock Paintings by protecting the it with high precision and fidelity. Thanks for the help of digital platform, we can disseminates the cultural heritage resources of Cangyuan Rock Paintings scientifically and effectively. In today's increasingly fierce media competition, there is no doubt that the birth of new media creates a younger dissemination atmosphere and a more explosive dissemination effect for this rare and

precious culture heritage. The application of digital technology, such as AR and VR, not only presents the unusual cultural landscape in front of the broad audience and makes them feel as if they were personally on the scene, but also is conductive to the protection and cultural inheritance of Cangyuan Rock Paintings.

References

1. Wei, M., Jing, S.: Study on the ideological implication of Cangyuan rock paintings. Chin. Newspaper Ind. (2013)
2. Ruimin, Z.: Digital protection of intangible cultural heritage abroad and its enlightenment—Concurrently discuss on digital protection strategy of Wuqiang new year pictures. Data Cult. Educ. (2015)
3. Zheng, J., Gao, Q.: Exploration and research on digital inheritance method of XiChou rock paintings in Wenshan Prefecture, Yunnan Province. Art Lit. Masses (2013)
4. Bai, T.: Analysis of the possibility of digitalization of murals in the Mogao Grottoes. J. Changjiang Univ. (Soc. Sci. Ed.) (2011)
5. Dan, C., Lu, R.: Analysis on the digital development of new year picture art in Zhuxian Town. Packag. Eng. **38**(6), 62–65 (2017)
6. Bo, Z.: Brief Analysis on Application of Cangyuan Rock Paintings in Two-Dimensional Animation. Dalian Polytechnic University (2014)
7. Bai, T.: A rustic opinion on the importance of digitization of murals in the Mogao Grottoes. J. Arts Educ. Res. (2011)

How to Construct and Apply Self-learning Support System of Art Curriculum in the New Media Environment

Zhiyuan Ma[1], Jingjing Guan[2(✉)], Eeljin Chae[1], and Di Wu[3]

[1] China Korea School of New Media,
Zhongnan University of Economics and Law, Wuhan 430072, China
18062111221@163.com, cinetree@dongseo.ac.kr
[2] Institute of Education Sciences, Wuhan University, Wuhan 430072, China
18971083933@163.com
[3] School of Education Science, Hubei University, Wuhan 430072, China
wudi5@ustc.edu

Abstract. In the new media environment, the art curriculum autonomous learning support system is based on the new media information technology to build an intelligent computer big data computing aided learning environment. The system cooperates with the appropriate teaching guidance method to promote art learners to become artistic talents with deep basic knowledge and active creativity. The integration between the new media technology and the autonomous learning support system not only brings about the innovation of learning methods, concepts and tools, but also puts forward new demands for the individualized service of the art curriculum self-learning. The paper proposes the process and method of the autonomous learning service model of art curriculum. This paper analyzes the self-learning service of art curriculum based on learning situation, and position information based on the reasoning, the reasoning of behavior patterns and the reasoning of association rules, constructing context-aware art learning data processing and reasoning process.

Keywords: New media · Art curriculum · Independent learning · Context awareness

1 Introduction

With the popularity of mobile terminals and the prevalence of learning social platforms, various educational methods help scholars to use the fragmented learning time after work, or systematically study various types of education. The concept of independent learning complements the characteristics of new media, network, intelligence, service and synergy, and is the starting and innovation point for constructing a new ecology of independent learning. The self-learning support system contains a new generation of information technology such as the Internet of Things, cloud computing, big data, and ubiquitous networks. It can help improve learners' flexibility, adaptability, and initiative to combine self-directed learning [1]. This paper attempt attempts to integrate new media and new technologies, and complete the construction of the self-learning support

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 128–133, 2020.
https://doi.org/10.1007/978-981-32-9244-4_17

system model for art curriculum through the calculation of context awareness data. It is hoped that the self-learning support system model of the art curriculum based on the Internet-based infrastructure and the new media environment will promote the deep integration of independent learning and new media.

In the era when new media is widely used, learning needs are everywhere, learning processes are happening all the time [2]. Learning scenes are no longer fixed to classrooms and schools, but become an educational ecosystem that connects schools, families, art galleries, libraries, and even virtual learning communities and public social platforms through the Internet. At the same time, the autonomous learning support system reflects the students' autonomy and cultivates students' subjective initiative, and believes that in the learning process, the knowledge acquisition is the result of the individual's active construction when acquiring information. The continuous innovation of learning methods, learning tools, and learning concepts has made knowledge acquisition more focused on active participation and collaborative construction, and the behavior of individual learning has gradually gained attention. The main purpose of this paper is to construct a learner-centered autonomous learning support system model and improve the learning effect of independent learning in art curriculum.

2 Methods

2.1 Model

The new technology covered by the current new media environment is more extensive and novel. It is not only a tool for information dissemination, but also subverts the traditional way of knowledge transformation. The concept of new media education has further brought about changes in education and learning models. This kind of change must be an innovation from the educational concept to the teaching method. Then it needs to integrate the new media with the new concepts, new behaviors and new models of teaching and learning, and build a new learning support service system. The basic connotation of self-learning education for art courses is based on the latest new media information technology to build an intelligent big data aided learning environment, using appropriate teaching methods to promote art learners to become artistic talents with deep basic knowledge and active creativity. It requires the latest computer technology such as internet of things technology, big data analysis and artificial intelligence to share a lot of tedious and mechanical repetitive work, so that teachers can better grasp the students' learning situation in a timely manner, and let students focus more on complex and valuable artistic learning and creative tasks. To achieve this goal, we need to fully grasp the learning characteristics and style of art learners, perceive their learning situations and interests, and provide suitable art learning resources and convenient interactive tools according to local conditions, intelligently improve the learning efficiency of each art learner.

Specifically, the process and method of the autonomous learning model for the art curriculum follow the four basic principles: learning resource matching, self-consistent teaching logic, rich learning experience, and timely feedback. It is represented by four targeted service components as shown in Fig. 1: the orientation of context awareness

and artistic learning needs; adaptive learning process; differentiated teaching evaluation; semantic-based art learning resource organization and reconstruction. In general, the following basic characteristics should be highlighted: a comprehensive perception of the learning situation and the learner's current artistic interests and needs; providing learners with personalized learning resources anytime and anywhere, mastering the learner's style and designing a variety of learning activities, differentiated learning diagnosis, learning advice and learning services.

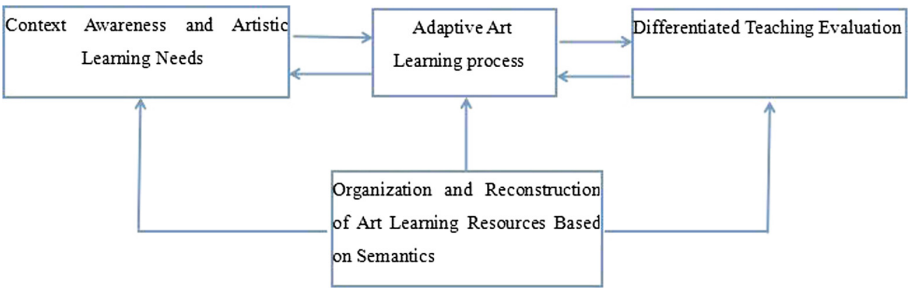


Fig. 1. Art course autonomous learning service model

2.2 Framework Construction

The application and development of Context awareness in educational technology is to adapt to the learning mode of autonomous learners' mobility and fragmentation. Context awareness technology stems from the study of ubiquitous computing, initially emphasizing computations that are integrated into the environment [3]. The ubiquitous network is the basis of context awareness. The ubiquitous network consists of a large number of non-traditional "smart items", which contains the various sensors and triggers, bring the perception of the environment to the smart device. Features such as virtual reality integration, multi-terminal access, seamless handover, and connected communities require computers (especially mobile computers) to adaptively change the organization and presentation of learning content through context awareness, providing push services for users [4]. Self-learning of art courses based on learning situations requires more knowledge of the situation in which learners are located through intelligent equipment, and perceives their specific artistic learning needs. For example, when a learner is in an art gallery, he may need to know relevant art knowledge based on GPS positioning and reasoning. If there is indoor positioning technology such as WIFI and RFID, it can be further precise whether he wants to know the art of painting or design. As shown in Fig. 2, context awareness features include not only time, location, user operations, but also a variety of sensor information. By comprehensively analyzing these characteristics, the user's contextualized information needs are generated and semantically represented. Analyze the actual learning needs of users as realistically as possible, semantically mark and extract the applicable contexts of learning resources, and correlate them to the behaviors of the targets, thus providing a more comprehensive and reliable basis for information push and filtering, and finally

the personality. The recommended information is displayed to the user in a differentiated manner [5].

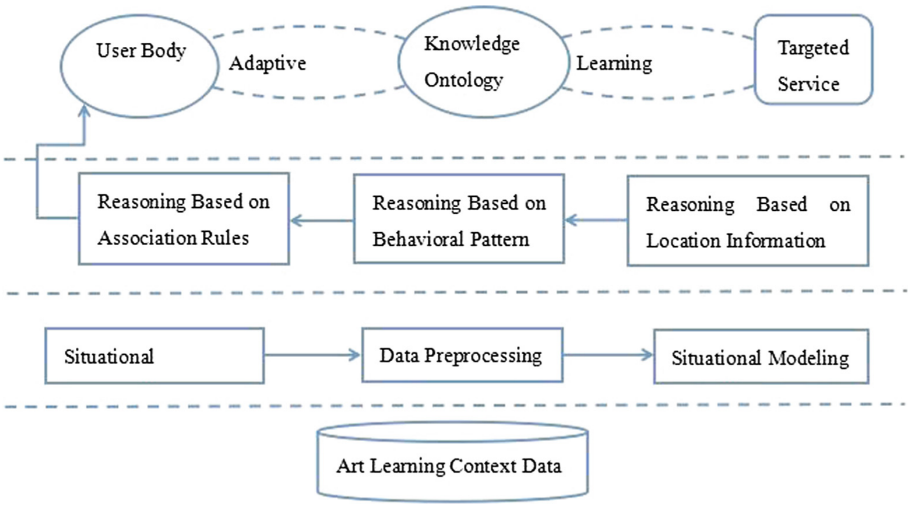


Fig. 2. Context awareness of art learning data processing and reasoning process

The focus of situational modeling is the construction of contextual contexts. The first thing to determine is which contextual elements need to be obtained. The basic contextual elements may include: the characters involved in the current context, resources used by the current context, learning habits and preferences of the user, current learning tasks, and the events, locations, times, perceptual data associated with the current context. Then based on the elements to reason, the process of reasoning can be divided into the following levels:

2.2.1 Inference Based on Location Information

Geographic location-based inference is the core of understanding individual's mobile behavior and learning state. It is very important for the recommendation application at a higher level, especially for the user's personal learning behavior. On the basis of the individualized and prominent location mining, combined with the existing artificially labeled sample, time and other context information for association rule analysis, the mapping rules of location type and learning state are summarized. The continuous location switching in the mobile scene have some limitations for the traditional behavior pattern recommendation on a single static location, so context learning is more concerned with continuous motion trajectory pattern mining. The definition of trajectory data is a serialized record of a set of pause locations and movement trajectories. The mapping of geographic information and semantic information is realized according to the movement trajectory and geographic information data, and the semantic knowledge and transition probability of the sampling points are introduced,

and the sequence mining algorithm is utilized. To analyze the semantic trajectory pattern and correlate learners' learning state changes.

2.2.2 Inference Based on Behavioral Patterns

User behavior pattern mining is the key to personalized recommendation of mobile users. It reveals the basic information of user learning rules and learning preferences, which is essential for improving recommendation efficiency. The situation of the learner and the real-time needs are closely related. In different situations, his recommendation needs will be very different. The effective analysis and utilization of mobile context data can better understand the user's current intentions and interests, and improve the recommendation system experience for mobile users.

2.2.3 Inference Based on Association Rules

As an important basis for carrying out targeted services, the association between user context and learning needs directly determines the quality of the personalized service and recommendation. Inference based on association rules is not only to discover the relationship between context and behavior, but also to reveal a strong connection between different contextual features. This task can be effectively accomplished through basic association rule mining techniques and extensions based on this. For example, the video of the art creation software that the learner watches in a certain situation indicates how the course content will be switched at the next moment. May be search relevant courseware or do the creative design exercise after-school. The learning system needs to promptly and accurately recommend the material content of the after-school contact, relying on the mining and discovery of association rules.

3 Discussions

The learning guided by the autonomous learning support system of the art curriculum in the new media environment can be adapted to most students. Based on big data computing and contextual awareness, this self-learning system has a targeted and differentiated adjustment of teaching, stimulating students' independent learning interests, mastering students' basic art theory and artistic creative ability. The promotion has been significantly helpful. Specific learning details and depth of understanding require students to rely solely on their own efforts to obtain. This way of teaching and learning leads to the knowledge being "superficially grasped", and it is better to learn and guide the personal interest matching without differentiation. This way of teaching and learning leads to the knowledge being "superficially grasped", and There is no personalized learning guide for personal interest matching.

4 Conclusion

Guided by the autonomous learning support system for art courses in the new media environment can be adapted to most students. This kind of guiding mode can adjust the teaching in a targeted and differentiated way based on big data calculation and

situational awareness, and stimulate students' interest in independent learning. It will greatly help students master the basic art theory and improve their artistic creative ability.

In most cases, the traditional unified teaching is to build a subject knowledge framework for students. The specific learning details and the depth of learning need to be obtained by students relying solely on their own efforts. With the development of new media technology, the auxiliary teaching of the self-study support system of the art curriculum, the time for students to understand the knowledge of art theory will be significantly shortened, and the ability to comprehensively create art will be effectively improved according to personal strengths. In turn, students will have a new experience in the practice of artistic creation, and further promote students' deeper exploration and study of art learning. Therefore, the teaching guided by the autonomous learning support system of art courses in the new media environment can be superior to the performance of traditional classroom learning.

Acknowledgments. This work was supported by Hubei Higher Education Teaching Research Project. "Construction of Advertising Training Teaching System Based on Industry Capacity Requirements" (No. 2016452); Hubei Humanities and Social Sciences Project "Research on the Reform of ESP in Applied Universities from the Perspective of Synergetics Theory" (No. 2017G009); "The Fundamental Research Funds for the Central Universities", Zhongnan University of Economics and Law (No. 2722019JCT042).

References

1. Yu, S., Chen, M.: Characteristics and trends of ubiquitous learning resources construction—taking the learning meta-resource model as an example. Mod. Distance Educ. Res. **6**, 14–22 (2011)
2. Holzinger, A., Motschnigpitrik, R.: Student-centered teaching meets new media: concept and case study. Int. Forum Educ. Technol. Soc. **10**, 160–172 (2002)
3. Premlatha, K.R., Geetha, T.V.: Learning content design and learner adaptation for adaptive e-learning environment. Artif. Intell. Rev. **13**, 1–23 (2015)
4. Li, W., Wang, W., Mo, T., Zhang, Z.: Review of situational computing research. J. Comput. Res. Dev. **2**, 542–552 (2015)
5. Shu, H., Wang, F., Cai, Y.: Construction of a new micro-group teaching model for group wisdom. Mod. Educ. Technol. **8**, 19–25 (2015)

Design of Online Multimedia Homework Management System Based on Cloud Platform

Si Fu[1(✉)] and Fei Hao[2]

[1] China-Korea Institute of New Media,
Zhongnan University of Economics and Law, Wuhan, China
miaancen@qq.com
[2] School of Computer Science, Shaanxi Normal University, Xi'an, China

Abstract. With the uprising of new media industry, homework of students majored in mass media becomes increasingly diversified. To better deal with the multi-media homework processing demand and achieve smoother communication between mass media teachers and students in term of latter's homework, this system makes some improvement to the existing homework review system that focuses on text work, including introduction of video/image homework processing and multi-media work exchange and sharing. An online multi-media homework management system is designed here on the basis of Aliyun platform. The system featuring comprehensive and powerful functions can provide better assistance for such teaching activities as lecturing of teachers and students engaged in mass media industry.

Keywords: Multimedia homework management · Browser/Server mode · Cloud platform

1 Introduction

As science and education keep advancing, network has gradually become a new platform of education [1–5]. In colleges and universities, network gets widely used as a critical auxiliary tool for education [6, 7]. Network teaching, a novel teaching mode, is advocated on an on-going basis on the other hand [8–10].

As one of the important links in teaching, homework has feedback and communication functions. After class, by the means of the internet social tools such as E-mail and QQ, the homework submission function is realized. However, this can only solve the problems such as simple homework delivery and submission, and it is impossible to make together the behaviors such as homework correction, homework feedback, and homework query statistics between teachers and students for further analysis and judgment. There is a break-up between each other, and it is necessary to build a platform for teachers and students to collaborate synchronously [11].

With emerging new media, the homework of mass media students is no longer monotonous in form; instead a variety of new forms including image, video and audio are employed in homework. The existing online homework management systems are mostly single-functional and can't meet the increasingly diversified homework file processing demand by the students majoring in mass media (such as its inability to deal

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 134–139, 2020.
https://doi.org/10.1007/978-981-32-9244-4_18

with the homework in image, video and audio forms). Therefore, an improved online homework management system is proposed here in this paper, which offers preliminary format check for the video and image homework in new media-related majors to facilitate the teachers' teaching and homework reviewing activities on the basis of those basic functions such as homework submission and reviewing.

2 Design of System Functional Modules

According to the actual requirements of the multimedia department, the system mainly consists of following functional modules: administrator module, teacher module, student module and exchange module.

2.1 Administrator Module

The administrator module is provided with the following functions:

System administration: on/off and properties settings of the system and regular system maintenance; Administrator administration: manage and maintain the information about the administrator user, such as the identity, category and limits of authority; Teacher administration: add/delete and maintain the teacher-related identity and authority limits information, and design corresponding class and course for specific teacher; Student administration: add/delete and maintain the identity and authority limits information about the student; Data administration: create relatively comprehensive data backup mechanism for the system to avoid any teaching accident arising from an occasional event and subsequent data loss. In addition, the system is also provided with a multi-media homework sharing and display platform. To ensure the working efficiency of the system, it is necessary to regularly clean the old homework saved in the system's database.

2.2 Teacher Module

In order to alleviate the teachers' workload in traditional assignment execution process and improve the teaching efficiency and quality, some rich and practical functions have been designed in the system for the teacher user on the basis of the major features. Those functions include:

Homework assignment: after successfully logging in the system, the teacher user can check his/her class and courses in present term. The teacher can add new homework for the lectured class as per corresponding course. The homework is mainly assigned in three forms: text, image and video. The teacher can specify both deadline and file attributes of the homework (such as word counts and text font and size for text work, and format and size of image or video work). The system will conduct preliminary review on the homework files submitted by the students as specified by the teacher for the homework and reject the non-conforming ones, and warn the students with pop-up window. Besides, based on the actual demand of multi-media major, the teacher may ask the students to submit the operational procedures around key nodes in finishing their homework when adding an assignment, especially when it is in image or

video form. When this item is checked, the students will be required to record the screens about their using the dedicated software to finish the homework with the plug-in provided by the system when approaching the critical steps in the assignment; whereas the system will upload and save the screen recording contents of the students for reference of teacher or later use in answering students' puzzles. This can help the teacher to better make sure whether the students finish the multi-media work as required and what problems they encounter during the process.

Homework management: the teacher user can make supplementary notes to the submitted assignment or terminate it when necessary after logging in the system. Besides, such user can also check the homework progress and submission by the students as well as the checked and reviewed or un-reviewed homework.

Homework statistics: the teacher can check the progress, overall completion and score distribution of the assignment among students of different classes in the system. Colleagues may carry out a lateral comparison and analysis of the homework finishing among different classes and the final analysis outcome can be exported from the system.

2.3 Student Module

The functions of student module mainly include:

Individual information management: after logging in the present system, the students can correct any error concerning their individual information and submit the corrected information. They can also set their telephone number and email address as well as enabling of short message or email notice in case of a new assignment.

Homework inquiry: the students can search and inquire the assignments released, unfinished or finished and submitted in the system, and check the outcome of their previous homework.

Homework submission: the students can submit files in specified format at any time before the deadline, and the homework will be successfully submitted after passing the preliminary format review by the system. When doing the homework, the students can also record the screens about their software operating process when approaching corresponding key links as required by the assignment.

Score inquiry: the students can check their score for a specific assignment and the scores for all the assignments under one course in a term after the teacher finishes the reviewing process.

Problem solving: the students may pose questions for their teachers in the system at any time before the course is finished. They can also insert the screen recording of their software operating process when submitting the questions so that the teacher can better find the causes behind the problems and provide accurate answers.

Mutual Evaluation Mode for Students.

The novelty of the design doesn't only lie in traditional teachers' evaluation system, a new evaluation mode is added, namely, a mutual evaluation mode for students. Big data and B/S mode of the Internet are adopted, Html, Css, Javascript, Jsp and Mysql technologies are adopted. After students upload their assignment and works, all of them will be displayed anonymously on the website, each student is required to grade at least ten students' work with remarks to successfully upload their work. The number of

times graded assignments are shown should be as consistent as possible. The first ten students uploading their assignments need to log in again and be evaluated by the system, or else, the uploading can't be counted as successful. 50% of each student's performance comes from the final scoring in mutual evaluation of students—the algorithm of big data is adopted, and calculation is conducted by taking 10 as the unit. In the end, the most frequently given score will be adopted, and the remaining 50% will come from the correction of teachers. This will be fairer for students' works.

2.4 Public Module

This module contains following primary functions:

Materials sharing: the teacher can selectively upload the courseware and materials about the course to the system for reference by the students, and upload, wholly or partly, the reviewed homework of certain course to the sharing zone in the system, and set whether the contents shared can be downloaded. When uploading and sharing are conducted, the system will automatically add watermarks of different forms to files of varying forms (such as official title of the school to video homework) so as to protect the copyright.

Exchange and interaction: both teachers and students can subscribe new contents prompt for public sharing zone so that once new works of proper authority are shared, they can comment on and mark them for mutual exchange and progress.

General information: this part serves to supplement and display the information about the system, including "about the system", "links", "campus landscape", and "feedback".

3 System Framework Design

The system adopts B/S (Browser/Server) three-layer architecture mode. Based on Web technology, this mode is a system platform model that decomposes the server part in traditional C/S (Client/Server) mode into one data server and one or more application Web servers in order to form a three-tier structure (user interface, service logic, and data/server) [10, 12]. The architecture is illustrated in Fig. 1:

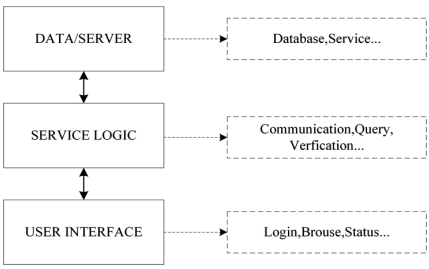


Fig. 1. The architecture of the system based on B/S mode

In accordance with module function design and general framework design, the present system is mainly composed of access webpage, portal page, workflow interaction control platform, homework reviewing service platform, public sharing and exchange platform, backstage management platform, and cloud computing service platform. The logic structure of the system is demonstrated in Fig. 2:

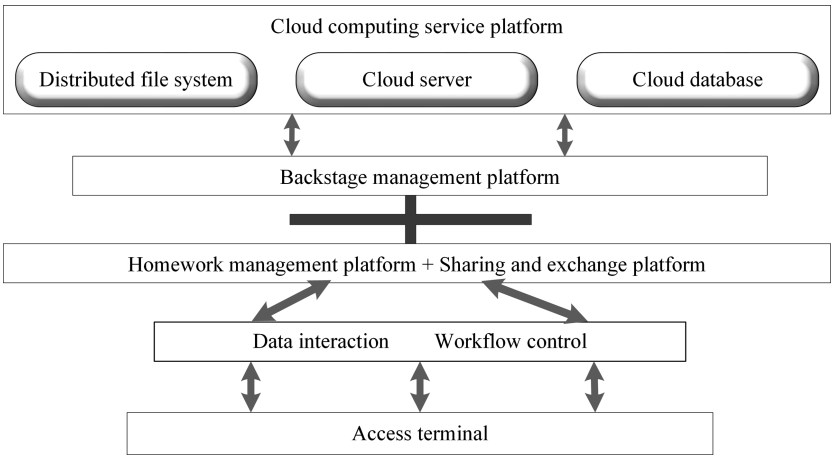


Fig. 2. The logical structure of the system based on cloud platform

The system employs HTML+CSS+JSP as the frontend interaction page development tool, JAVA language to achieve all sorts of operations in service logic layer, and Aliyun service as the database and server supplier. As a sort of business computing model, cloud computing assigns the computation load to the resource pool composed of a great number of servers so that the customers can obtain access to corresponding memory space in database, computing power and related information service as necessary [13]. The resource pool here is called "cloud" which is a large server cluster that is made up of computer server, storage server and broadband resource. The reason why Aliyun server is selected in this system is because the cloud host of Aliyun is built on the well-known Apsaras cloud-computing system that is widely used by such large enterprises as Taobao and Tmall due to its high reliability [14–16]. Besides, the system makes use of the distributed file system provided by Aliyun to efficiently complete the storage, invoking and editing tasks for homework files of a number of media forms in the system.

4 Summary

By carefully reviewing some existing online homework management systems, this paper proposes an improved and cloud platform-based online multi-media homework management system to meet the actual demand of new media majors. On the premise of containing homework submission, reviewing and inquiring functions in traditional

systems, this system integrates image and video conversion, downloading, saving and editing as well as forum communication functions to better ameliorate the teaching effect in new media majors and realize the "convenient and sharing" targets of education.

Acknowledgement. Supported by "the Fundamental Research Funds for the Central Universities", Zhongnan University of Economics and Law (No.2722019JCG079).

References

1. Anshari, M., Alas, Y., Guan, L.S.: Developing online learning resources: big data, social networks, and cloud computing to support pervasive knowledge. Educ. Inf. Technol. **21**(6), 1663–1677 (2016)
2. Bora, U.J., Ahmed, M.: E-learning using cloud computing. Int. J. Sci. Mod. Eng. **1**(2), 9–12 (2013)
3. Coronel, C., Morris, S.: Database Systems: Design, Implementation & Management (2016)
4. Chang, V., Gütl, C., Ebner, M.: Trends and opportunities in online learning, MOOCs, and cloud-based tools. In: Second Handbook of Information Technology in Primary and Secondary Education (2018)
5. Chen, S., Lin, M., Zhang, H.: Research of mobile learning system based on cloud computing. In: 2011 International Conference on e-Education, Entertainment and e-Management (ICEEE), pp. 121–123. IEEE (2013)
6. Dahdouh, K., Dakkak, A., Oughdir, L., et al.: Big data for online learning systems. Educ. Inf. Technol. 1–18 (2018)
7. Daradoumis, T., Bassi, R., Xhafa, F., et al.: A review on massive e-learning (MOOC) design, delivery and assessment. In: 2013 Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), pp. 208–213. IEEE (2013)
8. Huang, R.: Development of a cloud-based network teaching platform. Int. J. Emerg. Technol. Learn. (iJET) **13**(04), 176–186 (2018)
9. Jeong, J.S., Kim, M., Yoo, K.H.: A content oriented smart education system based on cloud computing. Int. J. Multimedia Ubiquitous Eng. **8**(6), 313–328 (2013)
10. Laudon, K.C., Laudon, J.P.: Management Information System (2016)
11. Masud, M.A.H., Huang, X.: An e-learning system architecture based on cloud computing. System **10**(11), 255–259 (2012)
12. Mehlenbacher, B., Kelly, A.R., Kampe, C., et al.: Instructional design for online learning environments and the problem of collaboration in the cloud. J. Tech. Writing Commun. **48** (2), 199–221 (2018)
13. Pocatilu, P., Alecu, F., Vetrici, M.: Using cloud computing for E-learning systems. In: Proceedings of the 8th WSEAS International Conference on Data Networks, Communications, Computers, pp. 54–59. World Scientific and Engineering Academy and Society (WSEAS) (2009)
14. Sultan, N.: Cloud computing for education: a new dawn? Int. J. Inf. Manage. **30**(2), 109–116 (2010)
15. Sun, G., Cui, T., Yong, J., et al.: MLaaS: a cloud-based system for delivering adaptive micro learning in mobile MOOC learning. IEEE Trans. Serv. Comput. **11**(2), 292–305 (2018)
16. Then, M., Ianniello, B.R., Vu, D.B., et al.: A competence-based course authoring concept for learning platforms with legacy assignment tools. GSTF J. Comput. (JoC) **4**(4) (2018)

On the Evaluation of Propagation Force of New Media for Government Affairs Based on the Theory of Information Acceptance Technology

Yuzhi Dong(✉)

China and South Korea School of New Media,
Zhongnan University of Economics and Law, 182# Nanhu Avenue, East Lake High-tech Development Zone, Wuhan 430073, People's Republic of China
dyz3122@126.com

Abstract. The development of Internet technology and the wide application of new media, such as the Weibo, WeChat or some kinds of new Apps, promoted the construction of new media for government affairs. The decentralized new media for government affairs has weakened the propagation force and influence, therefore, based on the theory of information acceptance, resource integration and cluster communication are the big trends of new media for government affairs. A set of system, which combine with the quantitative and qualitative evaluation, the content supply and user experience assessment, the communication function and efficiency evaluation, can be established to evaluate the clustering propagation force of the new media for government affairs may adjust the communication strategies, accelerate the communication efficiency, enhance the government credibility and promote the construction of service-oriented government.

Keywords: Information adoption technology · New media for government affairs · Cluster communication

With the rapid development of Mobile Internet Technology, the new media forms such as Weibo, WeChat and some news Apps, has changed the media landscape. Micro media has become an important carrier for people to communicate and obtain information. It is reported by the 42nd China's Internet Development Survey Report, the scale of Internet users in China reached 802 million, with 29.68 million new Internet users, a growth rate of 3.8%, and the Internet penetration rate reached 57.7%. Such a large scale of Internet users use the new media to obtain information, meanwhile, they constantly improve the propagation force of it, and promote the progress of our administrative information disclosure, forming a unique Chinese micro-era landscape. Innovating the evaluation system of new media communication of government affairs. There have been many empirical studies on the evaluation index design of government

Dong Yuzhi, female, doctor degree, professor of New Media Institute of China and South Korea, Zhongnan University of Economics and Law, majoring in new media and film cultural communication.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 140–146, 2020.
https://doi.org/10.1007/978-981-32-9244-4_19

new media performance and development level, but the development trend of government new media cluster platform has injected new content and changes into new media for government affairs. Yu Guoming believes that the propagation force is a kind of soft power, and puts forward that "the propagation force is the most important soft power that affects the society, and the most important power that affects people's identities of feelings, status and value [1]."

The information technology of government is mainly composed of network information technology (internet website, E-mail), telephone and communication technology (mobile phone, videophone, teleconference), multimedia technology (touch screen, digital TV) and security technology (smart card) [2]. There are three major influencing factors for government to adopt information technology: technology, organization and external environment. The implementation process of government information technology adoption mainly involves key factors such as organizational process change, organizational structure change and organizational staff. However, the effects are related to the relationship between the adoption of government information technology and the efficiency of public sectors, the adoption of government information technology and government trust, and the adoption of government information technology and public participation [3].

1 Combination of Qualitative and Quantitative Evaluation

Qualitative evaluation is a process of making judgments on the system risk status based on non-quantitative data such as researchers' knowledge, experience, historical lessons, policy trend and special variations. It mainly takes the case study of the in-depth interviews with the investigation objects as the basic data, and then, through an analysis framework of theoretical derivation, codes and sorts out the data, then draws the investigation conclusion. Quantitative evaluation is to use mathematical methods to collect and deal with the data and make value judgments of quantitative results for evaluation objects. It is objective, standard, accurate, quantitative and simple, but it couldn't evaluate all the factors appropriately in practical evaluation.

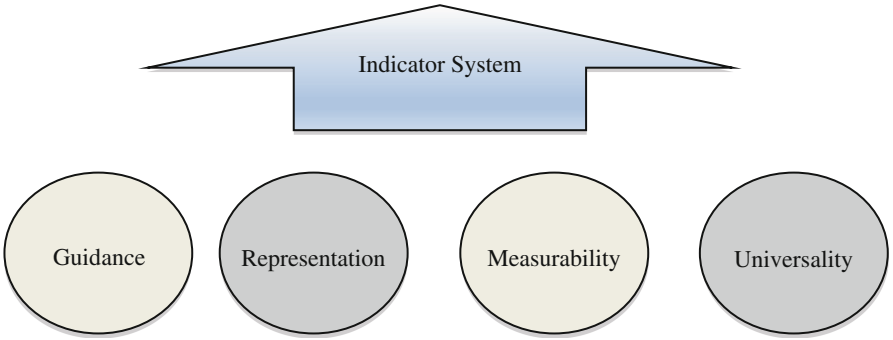
Qualitative evaluation does not use mathematical methods, but makes value judgment directly on the basis of the observation and analysis of real state or literature of the evaluation object. Qualitative evaluation emphasizes observation, analysis, induction and description. However, it sometimes makes the evaluation results vague and general, for its elasticity, which makes it difficult to grasp accurately. When we design the evaluation index system of the propagation force of local government new media clusters, we should adopt a qualitative or quantitative evaluation method according to the characteristics of specific indicators. Among them, qualitative research mainly adopts in-depth interviews, while quantitative research adopts three methods: online questionnaire survey, random sampling survey and face-to-face interview survey.

2 Combination of the Content Supply and User Experience Assessment

For the new media of government affairs, the evaluation of propagation force needs to return to the content level as the media itself, and analyze the content in different regions and departments. Through the evaluation of content relevance, usefulness, affinity and other aspects, to evaluate its influence and dissemination. At the same time, it is necessary to fully consider the difficulties and imbalance caused by external and internal factors, so as to make sure the rationality of content in the process of building the evaluation system. It should establish from the perspective of government new media constructors, examine the content supply, and based on the users' demands, enrich multiple user experience indicators, and guide the new media cluster to achieve a balance between service supply and demand satisfaction.

Current affairs of new media lacks user experience design thought during the process of construction, does not realize the core principle of "user-centered" in the content and function design, still adheres to the traditional propaganda thinking and one-way communication concept. From the change of audience concept in the new media era, the design of content and function should give priority to users. In the initial stage, users' needs should be fully analyzed and understood. For routine maintenance, it should be pay more attention on user's feedback, make adjustment timely, and analyze user's specific behavior and psychology, so as to improve and make plans (Table 1).

Table 1. Internal elements of the overall framework



The internal elements of the overall framework are mainly reflected in the following aspects: Firstly, government documents and laws and regulations. They are the summary and planning of the development of new media of government affairs in a certain period, which guides the direction of the development and provides a system and legal environment for its development. Secondly, the elements of government service. The application of new media of government affairs is the tool and path of self-innovation for government transformation to adapt to the "Internet +" era. Thirdly, the application

level of computer and communication technology. It is the key to the development of new media of government affairs. However, under the influence of different aspects such as economy, location, talent structure and technology investment, the application level of computer and communication technology of different local governments varies greatly. It has to do with local governments' adoption of new technologies and innovation. Fourthly, the degree of social demand. From the perspective of social needs, government new media is conducive to the extension of government functions. Fifthly, the individual level. The self-efficacy of mobile government affairs refers to the self-judgment of knowledge and skills needed in the use of mobile government based on the past knowledge, skills and experience.

Relevant studies have proved that the main reason for the low utilization rate of electronic services is that citizens generally lack the basic knowledge and skills to use online services. This shows that only when the public has a certain level of knowledge and skills, will they have the willingness to adopt. In addition, studies in related fields have found that trust affects users' attitudes.

3 Combination of Communication Function Assessment and Effectiveness Assessment

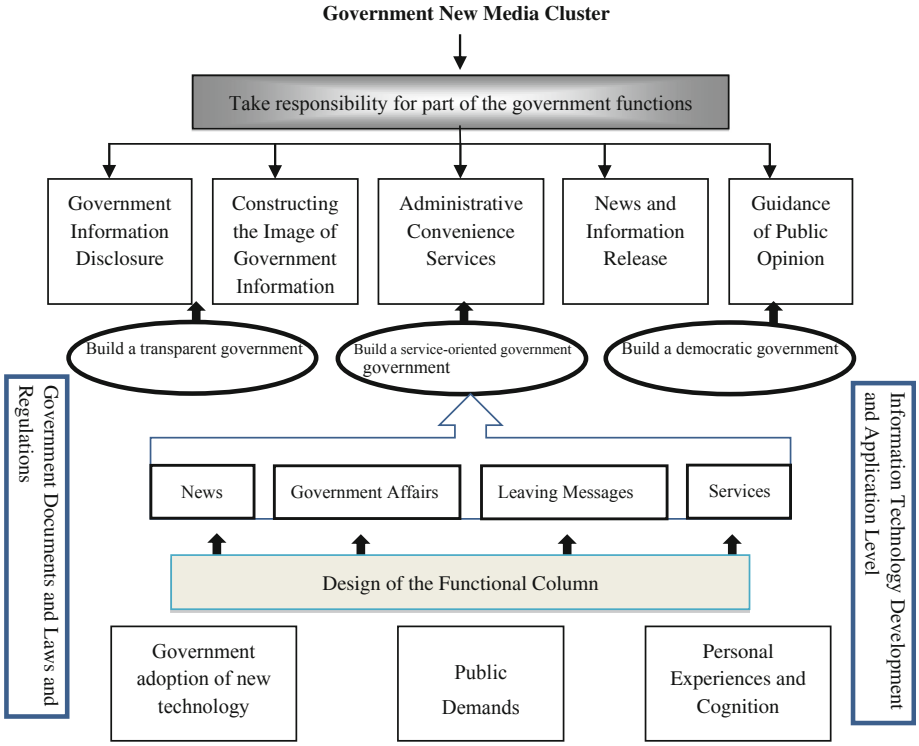
According to the interaction logic of the above factors and the communication function and effectiveness as the evaluation basis, the framework model of the evaluation system for the propagation force of local government new media cluster is shown in Table 2.

After the establishment of the evaluation principles and framework model of the evaluation system of the propagation force of government new media cluster, it is necessary to establish a scientific evaluation procedure to facilitate the standardization of evaluation: firstly, to determine the evaluation indicators. Focusing on the six aspects of government new media cluster: the information quality, the information disclosure, cluster degree, spread, interactive communication and public service, according to the existing function content of government new media cluster and the policy documents and the requirement of laws and regulations, the propagation force of the government new media should be considered, and the specific target content should be transformed into evaluation index to assess the strength of the government new media cluster from the intensity, speed, depth, breadth and service ability of the information, and thus reflects the dissemination ability of the main body of the new media cluster of government affairs and its impact on the audience's cognition and acceptance.

The second one is to determine the evaluation criteria. After defining the evaluation index of the propagation force of government new media cluster, the highest score in the evaluation index system should be determined according to the actual environment of the development of government new media cluster and the common needs of the audience, so as to evaluate the actual evaluation result of government new media cluster (Table 3).

The third is to determine the weight of index. Each index is weighted to determine the hierarchical evaluation system and form a general table of the evaluation index system. In any evaluation process, the determination of weight is prominent, because it reflects the status or role of each index in the evaluation, and directly affects the

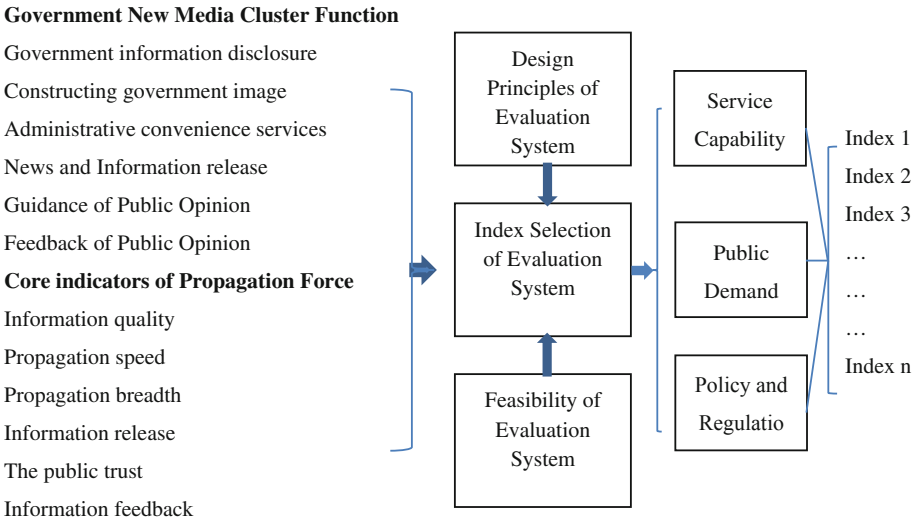
Table 2. The logic of the constituent factors in the evaluation framework of the propagation force of government new media cluster



evaluation results. It is also an indispensable part to determine the weights of the index in the construction of the evaluation index system of the propagation force of government new media clusters. At present, the main methods to determine the weight of index are subjective judgment method, expert consultation method, correlation coefficient method and factor analysis method. The expert consultation method usually gives the weight based on the experts' experience, which can reflect the actual situation to a certain extent, and the evaluation result is more realistic. Therefore, the weight determination in this study is mainly based on the consultation by experts, according to the theoretical research results, the development trend and the experience gained from the investigation and research of government new media stations, the weight of each index is determined by the analytic hierarchy process.

It is an innovation of the evaluation system of new media of government affairs to start with the evaluation of the mass propagation force. Therefore, in the face of the new situation, reflecting the current situation and guiding the development are the core evaluation of the propagation force of the government new media cluster. The innovation of the evaluation system can play a better role in guiding, promoting and promoting the evaluation, and provide practical basis for improving office efficiency

Table 3. The framework model of the evaluation system of local government new media cluster propagation force



and realizing smart government. Information technology adoption belongs to the category of organizational public management. Governments, enterprises, institutions and other organizations can better carry out information construction under the guidance of this theory. From information release tools to governance enabling, the technological revolution has become an important factor in the evolutionary trajectory of governmental transformation. From the perspective of the development level of government new media in different regions, the government new media cluster is only regarded as the medium of information release, and it does not realize the effective integration of user thinking, technological thinking and government information resources.

Therefore, in the high-dimensional Internet era, the establishment of the evaluation system of local government's new media cluster's propagation force is conducive to the scientific evaluation and ranking according to a set of standard and normative systems, thus reflecting the overall development of local and departmental government's new media clusters, in order to better use the evaluation to guide and promote the development of the new media for local government.

Acknowledgement. This paper belongs to the phased achievement of the Central Basic Research Business Project of Zhongnan University of Economics and Law, "Discourse Expression and Value Construction of Network Community Communication" (121#31721912102).

References

1. Nan, Z.: Study on TV promoting propagation force leveraging new media—the "Phoenix New Media" as an example. Master's thesis, Soochow University (2014)
2. Tan, H., Meng, Q., Zhang, N.: How information technology is applied to the government operation, a case of municipal government online administrative services construction in city. J. Sociol. Stud. **6**, 73–98 (2015)
3. Mei, L., Yi, Z., Wei, Z., Yi, Y.: A literature review on government adoption of information technology. E-Government **9**, 70–78 (2016)

Research on the Development Strategy of News Channel in the Context of Media Fusion

Shuang Zheng[1(✉)] and Fei Hao[2]

[1] New Media Institute of China and South Korea,
Zhongnan University of Economics and Law, Wuhan 430073, China
zssatine@hotmail.com
[2] School of Computer Science, Shaanxi Normal University,
Xi'an 710119, China
fhao@snnu.edu.cn

Abstract. With the rapid development of media integration in China, the traditional radio and television news channels are facing various opportunities and challenges, and the inherent mode of survival and development is becoming more and more unsuitable for the development trend of media integration. Therefore, how to strengthen the communication power, guiding power, influence and credibility of news channels in the context of media integration has become an urgent problem for radio stations to solve. This paper discusses the feasible development strategies from the perspective of news editing process, news broadcast development, the promotion of news anchorman's professional accomplishment and major news events.

Keywords: Media fusion · News editing · Live news · News anchor

With the rapid development of science and technology and the diversified demand of the market, various media show multi-function integration in the process of continuous development and gradually experienced the trend from competition to mutual reference and then to mutual integration. Moreover, the mass media, driven by the individuation of the audience's needs, speed up the transformation of the transformation to find a new mode of innovation. At the same time, with the rapid development of the technology of terminal display, information transmission and processing and intelligent sensing, the media fusion has obtained the technical guarantee. Under this background, the traditional mainstream media have been constantly exploring new development models, in which the TV news channels that shoulder the work and mission of the Party's news public opinion, how to further enhance the transmission power, influence, credibility and guiding power of news programs in the context of new media has become the media workers need to consider the issue.

1 Innovation of News Collection and Editing Process

Whether the content production of news channels has competitive advantage under the background of media integration depends on how quickly the editing process can adapt to the new media mechanism and effectively combine with the traditional editing

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 147–151, 2020.
https://doi.org/10.1007/978-981-32-9244-4_20

mechanism. How to break the old mode and establish a new all-media news editing mode is the first step of news channel development.

The process of news collection and editing begins with news gathering, which is different from the traditional news gathering mechanism. It is very imperative to integrate journalists' ranks in the process of new media and traditional media integration. Different from the traditional reporter interview mode, under the new mechanism, the reporter needs to have the processing ability of text, picture, audio and video at the same time. Too many manuscripts also decide that the reporter should have the media thinking when dealing with the news material. According to different communication platform customizing different manuscript content, the reporter need to be written for a classification of the same material.

For the news material sent back, the channel needs a powerful processing platform to aggregate and distribute a lot of content, and then through a platform to collect and integrate the news gathering clues from various channels, and utilize the platform to receive real-time live broadcast and form clues to catch the chain in time. Linkage "two micro-end" to realize the automatic set push, realize the free sharing of the platform, real-time interaction, the integration of editing and publishing search.

It is imperative to develop effective mechanism to coordinate pre-interview and post-editing. There are different requirements for content editing based on different distribution platforms, either focusing on in-depth analysis, or focusing on short and short reports, as well as on fragmentation and high frequency propagation, with different emphasis on different platforms. Later editorial departments need to communicate with reporters in time, that requires a new effective mechanism and a unified command department. Later editors set up groups according to the platform. After the contents of the interviews were uploaded through the platform, each group carried out the material download editing and final release according to the dissemination platform.

The innovation of news editing is gradual and requires constant adjustment and upgrading according to the birth and change of the new media. Only when the content meets the needs of all kinds of media can the influence and dissemination of the channel be enhanced.

2 Live Broadcast of Cross-media Cooperation

Live broadcast has always been the advantage and core competitive power of news TV channel. It attracts TV users by analyzing domestic and foreign hot spot events quickly and deeply and attracting hot spot tracking. However, in the trend of media convergence, this advantage is challenged by the new media. How to optimize the live coverage with the help of the characteristics of the media and consolidate the advantages of direct broadcasting to strengthen the influence and guiding power of the channel is the focus of the current development of the news channel.

Timeliness is still the key to the success or failure of live programs, especially news emergencies. The channel that first reports will take the lead. Therefore, the channel should take advantage of the advantages of new media to make efficient and reasonable reorganization of the live broadcasting department in order to increase the time limit of reporting. In the link of obtaining clues, in addition to the traditional hotline, we should

open a network platform in time, or we can also set up Weibo's news client, the channel's official WeChat feed platform. At the same time, the voice clue can be received and the network platform can be opened in a wide range to increase the video clue receiving terminal and the first scene can be seen in a quick and intuitive manner. In the link of live connection and video transmission, the channel should purchase new recording equipment and add new media broadcasting equipment to the original satellite vehicle to speed up live video transmission and ensure video quality. What we need to consider in the later stage of broadcast is not only the program form of TV platform, but also the diversity of forms, including "two micro-end", short video software and other platforms, so as to maintain high quality and efficient live news broadcast.

The selection of news channel live broadcast content is basically in the public sudden events or major public events or conferences, in which the major public emergencies have their natural audience cohesion. Under the trend of media convergence, news channels can make use of the characteristics of new media, expand the audience and expand their influence in major events. For major events or conferences, because the audience groups are relatively fixed, the channel want to expand the audience on the new media platform, it must find a unique perspective of live broadcast and find alternative interview points according to the news demands of new media users.

Combined with the new media platform, we shall to do a good job before the broadcast of live programs planning publicity and reprocessing playback after broadcast. While the live broadcast program release platform is increasing, it is rare to form a linkage. At present, the number of Internet users is close to 800 million in China, such a large number of users means that good publicity on Internet platforms will maximize the impact of programs. Prior to the broadcast of the live program, it is imperative to produced promotional manuscripts and short videos suitable for the new media platform to promote the program well in advance. After the program is broadcast, the production department can replay the program clip and produce fine products and put it on various media platforms to continuously increase the influence.

3 The Improvement of the Professional Literacy of News Anchor

Under the trend of media integration, the news program broadcasting platform is growing day by day, the demand of target audience is becoming more and more diverse, and the updating of technology and the usage habits of media users make every link of news program change. Among them, the form of direct broadcast is becoming more and more regular, and the ability of direct broadcast directly determines the effect and quality of this kind of program. Moreover, the convenience of communication makes the information source is no longer limited to the official platform anchorman, everyone can be the information publisher, in response to the news events voice, which is more stringent to the ability of the news anchor requirements. Based on various challenges and the trend of media convergence, anchorman teams need to further enhance their professional literacy to adapt to the development of the new era.

In order to cope with the trend of media convergence, news anchor's primary task is to improve the Internet thinking and understand the characteristics of new media to update the interactive hot spots in real time. In addition, News anchors shall analyze news events from the perspective of news recipients in order to enhance the "user" experience and analyze the current network users' psychology and thoughts, so as to do the right value guidance and news-oriented in a more easily understood and accepted manner.

Under the trend of normalizing live broadcasting, news anchor needs to have the ability to carry out news broadcast on its own when it goes deep into the news front line, which requires professional training for the anchorman of the news front line from news event planning to the use of portable camera equipment, computer video editing capabilities and integrated training. Under the influence of new media Internet thinking, this kind of anchorman also needs to plan and edit news content according to the users of the communication platform, skillfully use the new technology equipment, so that the program can spread more quickly and effectively to improve the program competitiveness.

With the help of the new media platform, to create the channel "network celebrity" anchor and enhance the popularity and reputation of news anchor. The external image of news anchor often determines the image of the audience to the news channel. The new media platform, such as "two micro-end" short video app, is used to increase the exposure of excellent anchors. According to the characteristics of the platform, the excellent anchors can attract new media audiences according to their own charm and accumulate the social influence and credibility of the anchors to enhance the channel competitive advantage.

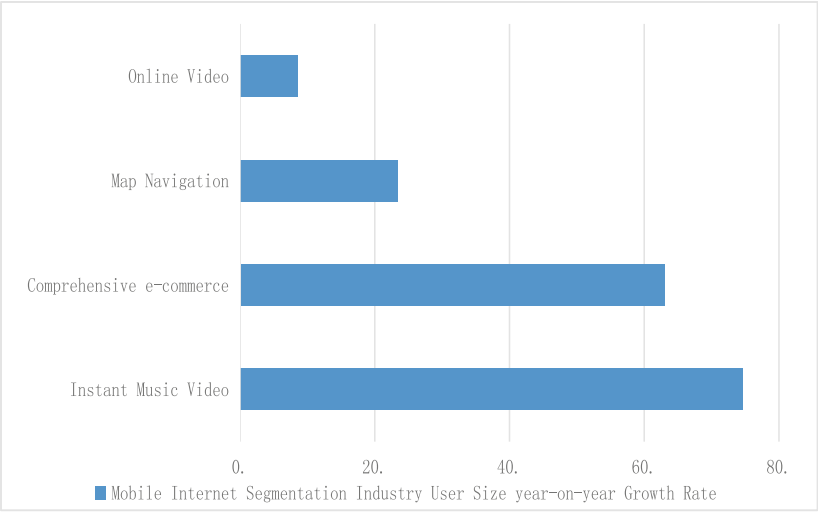
4 Upgrading the Major Theme News Reporting Model

Major theme news report is an important category of news for mainstream media to convey the voice of the party and the government and guide public opinion. It is also a link for all news channels to open their positions and show their competitiveness. Under the trend of media convergence, how to apply the advantages of new media to major media reports has become an important issue for mainstream media.

Using the characteristics of new media, carrying out cross-media live broadcast, covering the same subject, while ensuring the traditional advantages of live television, comprehensively excavating and reporting on other media from many aspects and from many angles, so as to deepen the audience's understanding of relevant events, supply related cognition, narrow the distance between traditional news TV program and audience, attract more multilevel audience and expand news influence.

Reasonable use of short video platform to do a good job in the major theme of the first stage of publicity preheating news reports landing short video platform to gain more audience and video views. According to the QuestMobile2018 annual data report, in December 2018, the total number of short video users per month increased by 33.1% year-on-year, ranking first in the top 10 industries. According to Douyin, the fastest growing short video application, its increase reached 557.6% year on the same month last year. The number of daily active users in China exceeded 250 million in 2018, and

the number of active users per month exceeded 500 million. Among them, the number of likes reached 2.83 billion in Beijing, flexible fragmentation and cooperation with mainstream media to increase the power and influence of news programs.



How to adapt to the current trend of media convergence and break through the limitations is an important proposition that media workers need to discuss all the time. At the same time, the mainstream media should be able to innovate, explore new development models and enhance the communication power, influence and credibility of news channels.

References

1. de Sola Pool, I.: Technologies of Freedom. Harvard University Press, Cambridge (1983)
2. Missouri Group. News reporting and writing. Bedford/S, America (2010)
3. Dailey, L., Demo, L., Spillman, M.: The convergence continuum: a model for studying collaboration between media newsrooms. Paper Presented at the Association for Education in Journalism and Mass Communication Conference, Kansas City, MO (2003)

Scrambler Based AES for Countermeasure Against Power Analysis Attacks

Young-Jin Kang[1], Ki-Hwan Kim[1], and HoonJae Lee[2(✉)]

[1] Department of Ubiquitous IT, Dongseo University, Busan 47011, Republic of Korea
rkddudwls55@gmail.com, ghksdl90@naver.com
[2] Division of Computer Engineering, Dongseo University, Busan 47011, Republic of Korea
hjlee@dongseo.ac.kr

Abstract. In modern society, cryptographic algorithms that guarantee the integrity of data have been used in various fields while being secured through various research. However, since the encryption algorithms are mostly used in the computer environment, there is a difference in the amount of leakage current depending on all calculation results. A power analysis attack is a method that utilizes a difference in leakage current and is effective in attack of a cryptographic algorithm. Therefore, we propose the structure of AES encryption algorithm to cope with power analysis attack by randomizing input using scrambler so that attacker cannot input arbitrary initial value and secret key.

Keywords: Power analysis attacks · Scramble · AES · Side-channel attacks

1 Introduction

The Advanced Encryption Standard (AES) [1] is a powerful encryption algorithm used to ensure the security of most communications, finance, and documents. The safety of AES is guaranteed by a mathematical field, so a common attack method requires a large amount of resources and time. However, side-channel attacks are used to attack physical information statistically instead of attacking cryptographic algorithms indiscriminately. Because physical information is varied by power, electromagnetic wave, temperature, sound, etc., attack methods are also diverse [2]. In particular, power analysis attack is a method widely used in side-channel attack because it measures leaked power to power source. At this time, the attacker enters arbitrary initial value and secret key into AES and tries attack. In this paper, we propose a method to protect the AES encryption algorithm by blocking the scrambling method of inputting arbitrary initial value and secret key for power analysis attack.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 152–157, 2020.
https://doi.org/10.1007/978-981-32-9244-4_21

2 Related Work

Power analysis attacks are known to be the most powerful of side-channel attacks, and known crypto algorithms have been shown to be vulnerable by power analysis attacks. As a result, many researchers say that a variety of studies are needed to counteract the power analysis attack [3].

Common countermeasures against power analysis attacks include inserting the defense logic at the hardware gate design stage and adding hardware logic or software coding according to the configuration method of the cryptographic module. According to the crypto algorithm, countermeasures against attack of power analysis attack should be separately designed.

The reason why the scrambler can be utilized as a countermeasure method of the power analysis attack is as follows. It is possible to countermeasure against the software structure without significantly altering the existing structure for the power analysis attacking the specific location of the encryption algorithm (AddRoundKey, S-Box, ShiftRows … etc.) [4].

A scrambler [5] is a method proposed for randomizing transmission data because clock synchronization recovery is difficult at the receiving end when the transmission data is "0" consecutively. At the receiving end of the communication equipment, it is converted back to a de-scrambler and then the data is input to the modem (Fig. 1).

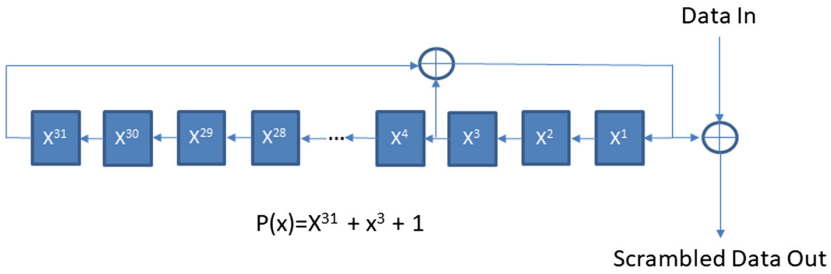


Fig. 1. Scrambler

3 Proposal Structure

The proposed scheme is a method of applying a scrambler to key and sub key values. The key length used in the AES is 128 bits, and the value is changed to a scrambler every time the key value is input in 1-bit unit and is input to the key schedule process. Based on AES 10 rounds (with a 128-bit key length), key randomization can be applied to number 11, 128 bit scrambles. The same applies to the receiving end.

Next, the plain text/round input value randomization is a method of applying a scrambler to a plain text value. The length of the plain text used in AES is unlimited, and each time a plain text value is input in 1-bit unit, the value is changed by a scrambler to XOR with the plain text/round input value. The same applies to the receiving end (Fig. 2).

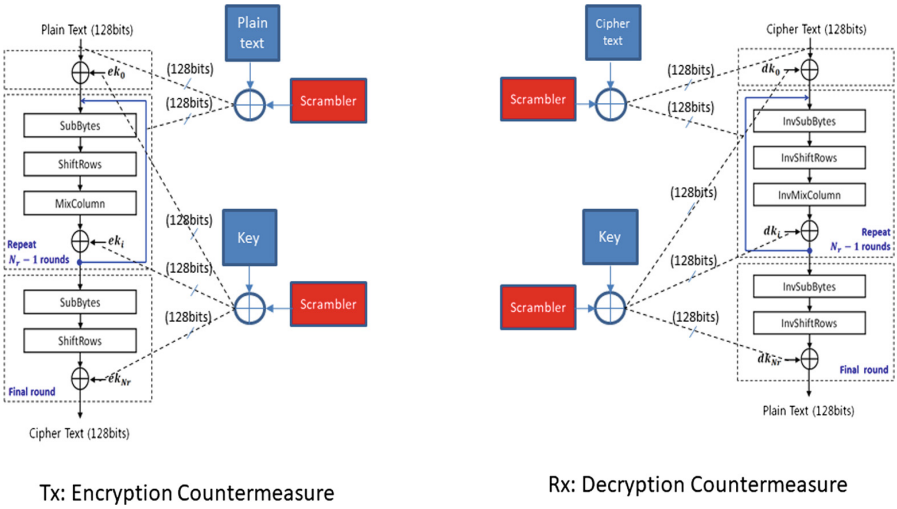


Fig. 2. AES scrambler structure

The scrambler generator uses a stream cipher structure that can generate a cycle of time without generating duplicate results. Stream ciphers can generate unique data sets at a frequency that is proportional to the memory size (N). AES uses 128-bit plain text and secret keys for each round's AddRoundKey operation. If the memory size of the scrambler is 8 or more, it can generate cycles of 128 bits or more. If the scrambler is activated before AddRoundKey is called, the start and end of the cycle cannot be distinguished. We changed the scrambler as shown in Fig. 3 so as not to affect the results of the existing AES calculation.

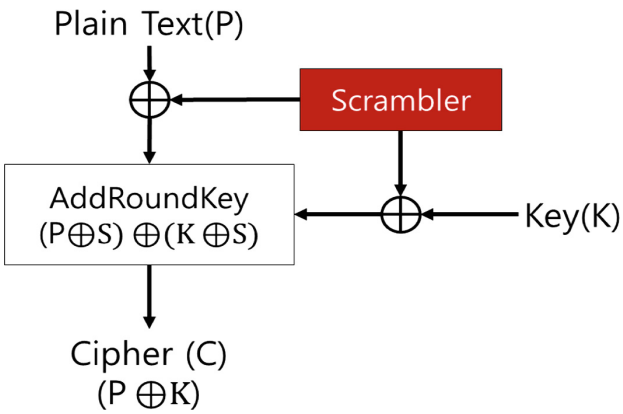


Fig. 3. A scrambler added to the AES AddRoundKey function.

As shown in Fig. 3, the scrambler naturally disappears in the AddRoundKey operation, and the operation result is the same as that of the existing AES. The equation is expressed as follows.

$$(P_i \oplus S_i) \oplus (K_i \oplus S_i) = (P_i \oplus K_i) \oplus (S_i \oplus S_i) = (P_i \oplus K_i) \oplus 0 = P_i \oplus K_i = C_i \quad (1)$$

3.1 Unable to Guess Based on Untrusted Initial Values

It is impossible to guess the correct key even if a good result is obtained in the consumed power signal by attempting to attack using the randomly generated plaintext and the secret key. The scrambler has a characteristic of a stream cipher, and the stream cipher has a property of not returning '0', so that the relationship between all specifiable secret key combinations (G) and the secret key combination (R) randomly modulated by the scrambler Can be defined as follows.

$$\mathrm{G} \oplus \mathrm{S} = \mathrm{R}, (\mathrm{S} \neq 0) \quad \therefore G \neq R \quad (2)$$

The power analysis attack is based on the fact that the secret key (K_t) can be known by replacing the power consumption signal (T_t) with the cipher text result (C_t) at the moment when the plaintext and the cipher text are made.

$$\mathrm{P_t} \oplus \mathrm{K_t} = \mathrm{C_t} = \mathrm{T_t} = \mathrm{T_t} \oplus \mathrm{P_t} = \mathrm{K_t} \quad (3)$$

That is, all specifiable secret key combinations correspond to a set of sequential trusted data that is not allowed to be duplicated, but due to the scrambler, there may be redundancy and become unreliable data.

$$\mathrm{P_t} \oplus (G_t \oplus \mathrm{S_t}) = \mathrm{P_t} \oplus \mathrm{R_t} = \mathrm{T_t} \neq \mathrm{P_t} \oplus \mathrm{K_t} \quad (4)$$

Therefore, even if the calculation is performed in consideration of all cases, the power consumption result cannot be correctly interpreted unless the interference of the scrambler is excluded.

3.2 Scrambler Influence Between Different Areas

Even if the initial values of the AES scrambler used in the transmitting side and the AES scrambler used in the receiving side are designed to be different from each other, the AES scrambler does not have any influence on the encryption/decryption result.

In the previous theorem, we confirmed that the scramble guarantees the computation result when it is used in each computation process. Therefore, even if the scrambled data on the transmission side and the scrambled data on the reception side differ from each other as shown in Fig. 4, it can be proved that no problem occurs in the result of the encryption/decryption.

If you attack the scrambler before attacking the AddRoundKey, you must attack the structure and initial value of the scrambler. However, the power analysis attack method

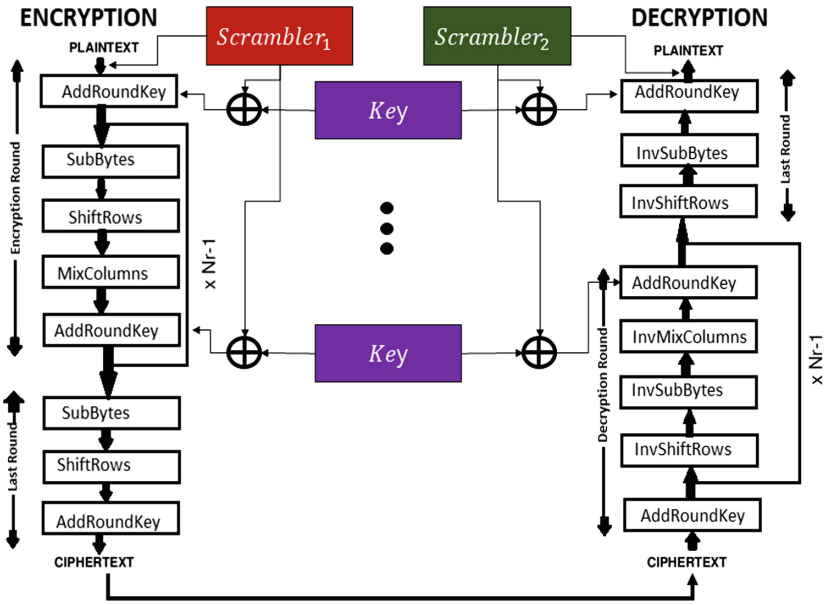


Fig. 4. Encryption/decryption issues using different scramblers.

is not to find the initial value directly but to combine all possible methods and select the one with the highest matching rate.

It is assumed that the initial value used for the scrambler is 2^8 or more. For example, if the scrambler period is set to the minimum value of 2^8, the attacker must perform and record all operations corresponding to $2^{136}(2^{128} * 2^8 = 2^{136})$.

4 Conclusion

In this paper, we propose a secure AES algorithm for power analysis attack. The power analysis attack on AES cryptographic algorithm proved that it is unsafe because many researches have been done on the point of time when XOR operation of plain text and secret key is performed or part of each function is attack point. However, when randomizing the key/sub key and plain text/round input values as described above, power analysis attack becomes difficult because randomization is performed for each round.

Acknowledgment. This work was supported by Institute for Information and Communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No.2018-0-00285) and also supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (grant number: NRF-2016R1D1A1B01011908).

References

1. AES cryptographic algorithm. http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf
2. Kocher, P.C., Jaffe, J., Jun, B.: Differential power analysis. In: Wiener, M. (ed.) Advances in Cryptology - CRYPTO 1999. Lecture Notes in Computer Science, vol. 1666, pp. 388–397. Springer, Heidelberg (1999)
3. Jaffe, J.: Introduction to differential power analysis. Presented at ECRYPT Summer School on Cryptographic Hardware, Side Channel and Fault Analysis, June 2006
4. Kang, Y.J., Kim, T.Y., Jo, J.B., Lee, H.J.: An experimental CPA attack for Arduino cryptographic module and analysis in software-based CPA countermeasure. Int. J. Secur. Appl. **8**(2), 261–270 (2014)
5. Joo, J.S.: A study on the performance of SS-SLM scheme according to the scrambling length. J. Inst. Electron. Inf. Eng. **55**(9), 19–24 (2018)

A Keyword Extraction Scheme from CQI Based on Graph Centrality

They Pheaktra[1], JongBeom Lim[2], JongHyuk Lee[3], and Joon-Min Gil[1(✉)]

[1] School of Information Technology Engineering, Daegu Catholic University, Gyeongsan-si, Gyeongsangbuk-do 38430, Korea
pheaktra97@gmail.com, jmgil@cu.ac.kr

[2] Department of Game and Multimedia Engineering, Korea Polytechnic University, Siheung-si, Gyeonggi-do 15073, Korea
jblim@kpu.ac.kr

[3] Department of Big Data Engineering, Daegu Catholic University, Gyeongsan-si, Gyeongsangbuk-do 38430, Korea
jonghyuk@cu.ac.kr

Abstract. Recently, most of the universities in Korea is doing a lecture evaluation survey every semester. The continuous quality improvement (CQI) report is one of the most popular lecture evaluation service systems, which able to summaries and analysis the mean of evaluation reports. Since 2016, education office allows CQI system to begin uploads and analysis the CQI report in all subjects. To improve the school and support to students, the school has to do a lecture evaluation after midterm and final exam every semester. The problems are the school getting so long to make the report on students lecture evaluation. In this paper, we propose a summary keywords extraction method form CQI and represented as graph tools based on centrality. We expected that this method can be efficiently extracted the most important relation keywords from huge CQI data of each lecture evaluations to summaries for the report.

Keywords: CQI · Keyword extraction · Graph · Centrality

1 Introduction

Universities or institutions in higher education try to improve the processes of providing their programs and services in all disciplines [1, 2]. Continuous quality improvement (CQI) is a management philosophy that institutions of education and research use to help increase operational efficiency, and improve clients' satisfaction [3, 4].

The main purpose of CQI is to improve the quality of the teaching-learning processes and curriculum at the university level [5]. A CQI report of universities includes

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2016R1D1A3B03933370 and NRF-2018R1D1A1B07045838).

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 158–163, 2020.
https://doi.org/10.1007/978-981-32-9244-4_22

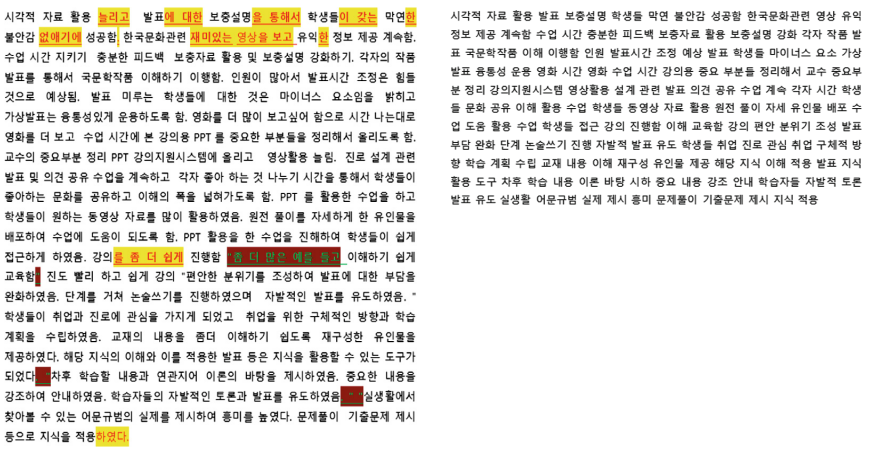
text data such as students' evaluation for lectures, self-evaluation of instructors for lectures, future improvement plans for lectures, and results of previous improvement plan [6].

A keyword is a significant word that represents the main contents of a document. In this regard, it is important to extract keywords from CQI report text data, which collectively form a kind of big data. By extracting keywords from CQI report text data, we can easily use that for various applications such as information retrieval [7, 8], data indexing [9, 10], clustering [11, 12], classification [13, 14], and content filtering [15, 16].

In this paper, we propose a summary keywords extraction scheme from CQI report based on centrality represented as graphs. The proposed keywords extraction scheme performs a pre-processing phase to deal with text data in Korean efficiently. Then, it maps word to word relationships based on word count and centrality. Since we provide a basic framework to extract keywords and construct word to word relationships as a graph, we can visualize the results and observe the summery of CQI report data.

2 Research Method

One of the difficult parts of processing text data is to trim unnecessary data. Thus, we design and conduct a pre-processing phase before performing the keywords extraction process as shown in Fig. 1. By performing the pre-processing phase, the size of CQI report data dwindled to about half of its original data volume.



(a) Before pre-processing (b) After pre-processing

Fig. 1. Before and after the pre-processing phase.

Figure 2 shows the flowchart of the proposed keywords extraction framework. The detailed steps for extracting keywords from CQI report data are as follows.

1. It remove unnecessary characters such as postpositional particle, symbols, numbers, one-letter characters, etc., from CQI reports.
2. For the collected data set, it performs the word count process. Then, it extracts top N words from the word count results.
3. It extracts keywords from the results of step 2 and maps word to word relationships based on graph-related operations.
4. In order to realize the relationships between the keywords, it constructs a graph from the results of word to word relationships.
5. Based on the constructed graph, it calculates centrality and weight values. Then it performs the normalization process with predefined scale.
6. It verifies the results based on F-score and visualizes the extracted keywords from CQI report data.

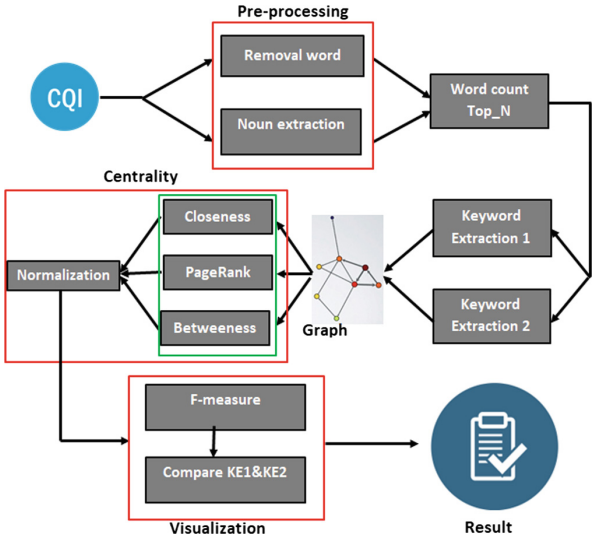


Fig. 2. Flowchart of the proposed keywords extraction framework.

For keywords extraction, we design two methods, that is, keyword extraction 1 (KE1) and keyword extraction 2 (KE2). In KE1, it maps the relationship between two words if the two words are adjacent. On the other hand, KE2 employs a sliding window method by extending relationship ranges. For calculating centrality, we use closeness centrality, page-rank centrality, and betweenness centrality.

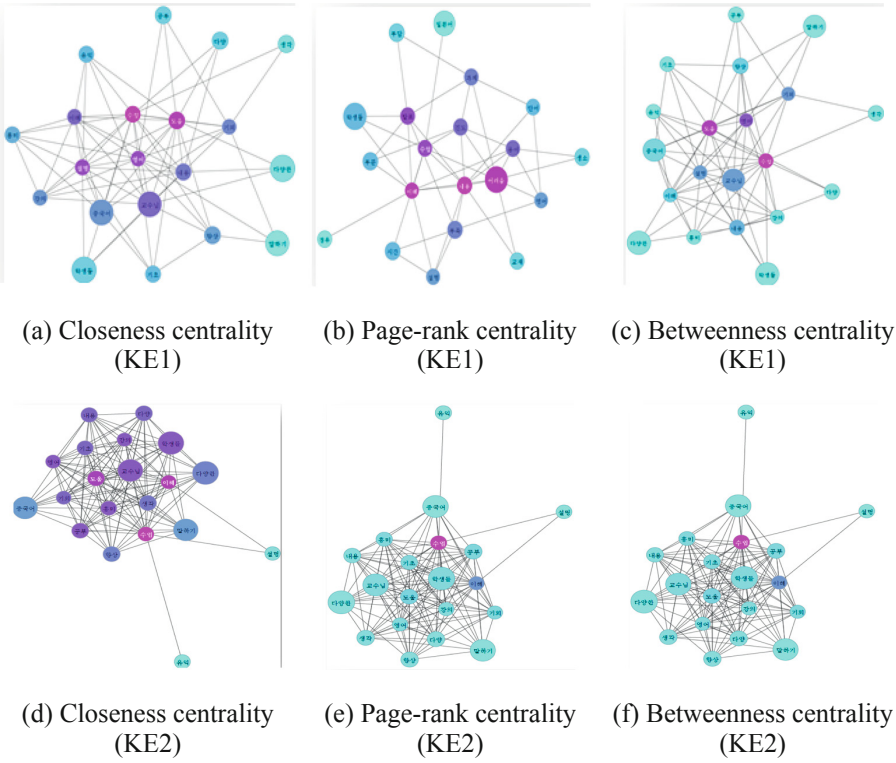
The closeness centrality is the reciprocal sum of the length of the shortest paths between a word to all other words in the graph. Then, it calculates the cost to spread words to all words in the document. The page-rank centrality is a variant of Katz centrality by taking eigenvector into account. The betweenness centrality of a vertex is the number of shortest paths that pass through the vertex. In the next section, we show evaluation results of our proposed summary keywords extraction scheme from CQI report based on centrality represented as graphs.

3 Experiment

Our proposed method aims to effectively extract keywords from CQI report data based on graphs. In this section, we provide experimental results of our proposed keywords extraction scheme.

We use 288 CQI report data of D University (2 semesters × 8 colleges × 6 domains × 3 years) in Korea from 2015 to 2017. The domains of the CQI reports include (1) good points, (2) weak points, (3) suggestions, (4) recommendations, (5) results of recommendations, and (6) self-evaluation of the instructor.

We develop keywords extraction scheme in Python for pre-processing, word count, and graph related modules. Figure 3 shows word and vertex relationships for CQI report data. The result graphs are generated by performing the pre-processing, word count, and top N phases through KE1 and KE2. Note that more frequent words are filled with a warm color.



(a) Closeness centrality (KE1) (b) Page-rank centrality (KE1) (c) Betweenness centrality (KE1)

(d) Closeness centrality (KE2) (e) Page-rank centrality (KE2) (f) Betweenness centrality (KE2)

Fig. 3. Experimental results.

The results indicate that comprehensive and complicated CQI report data can efficiently be processed by the proposed pre-processing phase. For frequent words, we can find the correlation between words in the document. The significance of words can

be calculated through the centrality measures. Hence, the proposed keywords extraction method is able to summarize the CQI data in a short period of time.

4 Conclusions

CQI is considered as an essential tool to help universities enhance their teaching-learning methods and approaches. In this paper, we propose a summary keywords extraction method form CQI and represented as graph tools based on centrality. With our proposed keywords extraction method, we can easily understand the meaning of huge CQI data. Our experiment reveals that the relationships between words are more meaningful than the basic top N words for the CQI report data. Our keywords extraction framework is able to find out the central connection between words and the strongest correlation word in the graph. Future work includes the verification of the results and designing more effective visualization methods with varying the number of the top N words.

References

1. Blouin, D., Tekian, A.: Accreditation of medical education programs: moving from student outcomes to continuous quality improvement measures. Acad. Med. **93**, 377–383 (2018)
2. Levy, G.D., Ronco, S.L.: How benchmarking and higher education came together. New Dir. Inst. Res. **2012**, 5–13 (2012)
3. Hogg, R.V., Hogg, M.C.: Continuous quality improvement in higher education. Int. Stat. Rev. Revue Internationale de Statistique **63**, 35–48 (1995)
4. Brown, J.F., Marshall, B.L.: Continuous quality improvement: an effective strategy for improvement of program outcomes in a higher education setting. Nurs. Educ. Perspect. **29**, 205–211 (2008)
5. Hedrick, J.S., Cottrell, S., Stark, D., Brownfield, E., Stoddard, H.A., Angle, S.M., Buckley, L.A., Clinch, C.R., Esposito, K., Krane, N.K., Park, V., Teal, C.R., Ferrari, N.D.: A Review of Continuous Quality Improvement Processes at Ten Medical Schools. Medical Science Educator (2019)
6. Lee, H.-C.: The Shift of the University Paradigm and Reform of the Korean University Systems (2005)
7. Berger, A., Lafferty, J.: Information retrieval as statistical translation. SIGIR Forum **51**, 219–226 (2017)
8. Fidel, R., Bruce, H., Pejtersen, A.M., Dumais, S., Grudin, J., Poltrock, S.: Collaborative information retrieval (CIR). New Rev. Inf. Behav. Res. **1**, 235–247 (2000)
9. Wang, J., Liu, W., Kumar, S., Chang, S.: Learning to hash for indexing big data—a survey. Proc. IEEE **104**, 34–57 (2016)
10. Gani, A., Siddiqa, A., Shamshirband, S., Hanum, F.: A survey on indexing techniques for big data: taxonomy and performance evaluation. Knowl. Inf. Syst. **46**, 241–284 (2016)
11. Mohebi, A., Aghabozorgi, S., Ying Wah, T., Herawan, T., Yahyapour, R.: Iterative big data clustering algorithms: a review. Softw. Pract. Exp. **46**, 107–129 (2016)
12. Aghabozorgi, S., Seyed Shirkhorshidi, A., Ying Wah, T.: Time-series clustering – a decade review. Inf. Syst. **53**, 16–38 (2015)

13. Shaikh, R.A., Adi, K., Logrippo, L.: A data classification method for inconsistency and incompleteness detection in access control policy sets. Int. J. Inf. Secur. **16**, 91–113 (2017)
14. Shaikh, R., Sasikumar, M.: Data classification for achieving security in cloud computing. Procedia Comput. Sci. **45**, 493–498 (2015)
15. Wang, Y., Ding, F.: Novel data filtering based parameter identification for multiple-input multiple-output systems using the auxiliary model. Automatica **71**, 308–313 (2016)
16. Deng, S., Yue, D., Zhou, A., Fu, X., Yang, L., Xue, Y.: Distributed content filtering algorithm based on data label and policy expression in active distribution networks. Neurocomputing **270**, 159–169 (2017)

Learning Influence Diagram Utility Function by Observing Behavior

Bai Lei(✉)

Department of Computer Science, Yunnan Normal University,
Kunming 650500, China
baileix@163.com

Abstract. This paper considers the task of learning utility functions of certain influence diagram based on the decision maker's past decisions. We assume that the influence diagram structure and the probability distribution it assigns to random events are known, so that we need only infer the utility function u for its. We also assume that the decision maker is rational. In particular, the decision maker's past decisions can be viewed as constraints on u. So, if we have a prior probability distribution $p(u)$ over u, we can then condition on these constraints to obtain u. In this paper, an approach for learning utility functions from decision maker's behavior was proposed. We also show that it is effective.

Keywords: Influence diagram · Learning utility function · Observing behavior

1 Introduction

Influence Diagram (ID) [1] has been widely applied to decision analysis and uncertain inference in many domains. When modeling a decision problem using the ID framework we need to specify both a qualitative part (represented by an acyclic directed graph, structure of ID) and a quantitative part (parameters of ID). The quantitative part is comprised of probabilities which represent the decision maker's uncertainty about the domain, and utility functions which represent the decision maker's preferences about the different outcomes of the decision problem.

Several approaches have been developed for the ID's structure learning (e.g., [2–4]). Also, the probabilities can be estimated using known techniques from machine learning and statistic. In this paper, we concentrate on the ID's utility function learning.

The acquisition of the utility functions is cognitively difficult and error prone; there is many elicitation techniques that often produce very different results when applied to the same person [5]. At present, some approaches eliciting the utility function based on a database of already elicited utility functions, some iterative refining the current utility function, and the others are based on observing the decision maker's behavior (e.g., [6, 7]).

Based on the assumption that the decision maker is rational, we proposed an approach for learning the ID's utility functions by observing behavior. In this paper, we also show that the utility function can be elicited effectively by using this approach.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 164–168, 2020.
https://doi.org/10.1007/978-981-32-9244-4_23

1.1 Influence Diagram

Influence diagrams (IDs) are acyclic graphs with three types of nodes: decision nodes, random nodes, and value nodes. The decision nodes represent decisions to make. The random nodes represent random quantities relevant to the decisions. Arcs into random nodes indicate probabilistic dependence and the dependence of a random node c upon its parents π_c is characterized by a conditional probability $P(c|\pi_c)$. The value nodes represent components of the utility function. Each value node v is associated with a utility function, a real valued function $f_v(\pi_c)$ of the parents π_c of v.

Here illustrates ID by using a decision-making problem of new products development in Fig. 1. We denote development-decision variable by D_1, price-decision variable by D_2, technology development cycle by X_1, states of development by X_2, the performance of new products by X_3, the condition of material supply by X_4, the marketing by X_5. V is the value node which indicates the decisional utility.

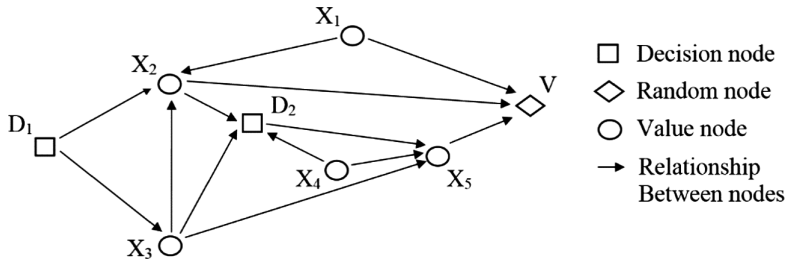


Fig. 1. An ID on decision-making of new products development

In the original definition of IDs [1], there is only one value node. In discussions on separability of utility function by Tatman and Shachter [8] allow multiple value nodes. We also allow multiple value nodes.

2 Learning Utility Function of ID by Observing Behavior

Any influence diagram can be expanded into a decision tree [9]. If we can observe the overall utility of each possible strategy (decision path), then we call it a fully observed strategy. That is, in the decision tree corresponding to the influence diagram, we observe that the decision maker of the other party selects each decision path from root to leaf and knows the overall utility of the decision path.

Assuming that every decision made by a decision maker comes from its rational thinking. If we observe that the decision maker has chosen a certain strategy s^*, then we can get the following inequality:

$$\forall_{s \in S} EU(s^*) \geq EU(s) \tag{1}$$

Where S is the set of all optional policies and $EU(x)$ represents the expected utility of policy x.

In a decision tree model, the decision-making process is from the root to leaf nodes through each node. The choice of path includes both artificial selection (at the decision node) and natural selection (at the random node). When "walking" to the decision node, according to the above inequality, the decision maker always chooses the child node with the largest utility value as the next passing node. The benefit of this decision should be equal to the utility value of the child node he selected. Therefore, for the decision node d, if a child node $\delta(\mathrm{n}) \in \mathrm{S(d)}$ was selected as the next target ($S(d)$ represents the child node set of node d), then the utility value of d is $\mathrm{U}_d = U_{\delta(n)}$.

For the random node r, it expresses a natural selection process. According to the conditional probability table (CPT) associated with it, the probability $P_{n'}$ of selecting any one of the child nodes $n' \in S(r)$ as the next target is known. So we can get the expected utility of random nodes r: $EU(r) = \sum_{n' \in S(r)} P_{n'} \cdot U_{n'}$. Where $U_{n'}$ is the benefit of choosing n' as the next target, that is the utility value of n'.

For the leaf node l, it is the last node in the decision path p, and its utility value U_l is the ultimate goal of the decision maker. We set the utility value of the leaf node as $U_l = \sum_{j=1}^{m} a_j v_j$, where v_j is the sub-utility, a_j is the weight of the v_j.

According to the above discussion, under the fully observed strategy, we can start from the leaf node and inversely derive the utility value of the nodes on all possible decision paths, then the utility value of each node in the decision tree can be obtained.

2.1 Learning Utility Value in Decision Tree

According to the above discussion, firstly, we expanded the influence diagram into decision tree. Under the promise of fully observed strategy, we can get the utility value of each node in the corresponding decision tree by the following algorithm:

Step 1: For the leaf node l, $U_l = \sum_{j=1}^{m} a_j v_j$, where m is the number of sub-utilities, v_j is the utility value of the *j-th* sub-utility, a_j is the weight of the v_j.

Step 2: For the non-leaf nodes n:

Step 2.1: If n is a random node, $U_n = \sum_{n' \in S(n)} P_{n'} \cdot U_{n'}$, where $S(n)$ is the set of child node of n, $P_{n'}$ is the probability of random events, $U_{n'}$ is the utility value of n' in the current decision path.

Step 2.2: If n is a decision node, $U_n = \max\left\{U_{n' \in S(n)}\right\}$.

Step 3: If the same decision path is observed m times, the utility values should be weighted average by weight.

2.2 Learning Utility Function of ID

Here we get the utility value U_i of each node in decision tree by the Algorithm presents in Sect. 2.1. At this point, we define the utility function F_v in the influence diagrams:

$$\mathrm{F}_v\left(\rho_v^j\right) = \frac{\sum_{i \in \rho_v^j} U_i}{\left\|\rho_v^j\right\|} \tag{2}$$

Where, ρ_v is the set of decision paths leading to node v in the corresponding decision tree, while ρ_v^j is the *j-th* decision path. $\left\|\rho_v^j\right\| = \|\rho_v\| \times$ the times of ρ_v^j appears in the decision tree.

3 Algorithm Verification

To verify the feasibility and effectively of our proposed algorithm in this paper, we constructed the influence diagram (ID_x) shown in Fig. 2, which missing utility functions.

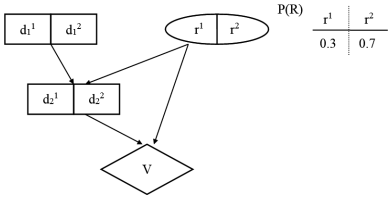


Fig. 2. An influence diagram which missing utility functions

The corresponding decision tree (T_x) shown in Fig. 3.

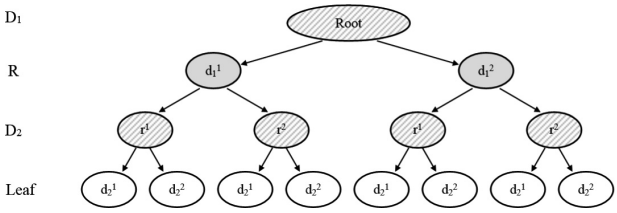


Fig. 3. The decision tree (T_x) corresponding to the influence diagram (ID_x)

Suppose we observed that the other decision-maker select each decision path from left to right, and the benefit of each decision path is 0, 20, 10, 60, 0, 50, 40, 30.

According to the above algorithm, the utility value of each node can be obtained, shown in Fig. 4.

So, the utility value of node v can be calculated:

$$\mathrm{F}_v\left(r^1, d_2^1\right) = \frac{0+20+0+50}{(2\times 2)\times 2} = 8.75, \quad \mathrm{F}_v\left(r^1, d_2^2\right) = \frac{20+20+50+50}{(2\times 2)\times 2} = 17.5,$$

$$\mathrm{F}_v\left(r^2, d_2^1\right) = \frac{10+60+40+50}{(2\times 2)\times 2} = 18.75, \quad \mathrm{F}_v\left(r^2, d_2^2\right) = \frac{60+60+30+40}{(2\times 2)\times 2} = 23.75.$$

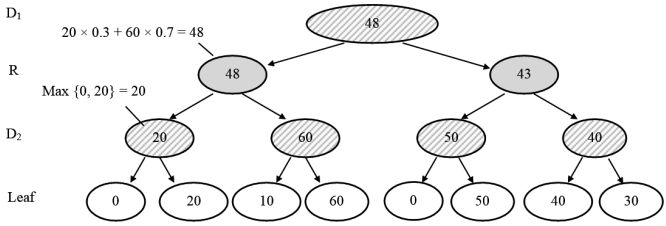


Fig. 4. The decision tree with utility value

4 Conclusions

In this paper, an ID utility function learning algorithm based on observing behavior is presented. Experimental results show that this algorithm is effective, and it requires a small number of samples. Although the more data observed is closer to obtaining more accurate calculation results, the method has higher accuracy than the previous methods for the case of less data. At the same time, for the case where the same decision path is observed multiple times, we can assign a higher weight to the new sample when performing the fusion with the old model, which is more accurate when performing the strategic predictions in a similar time.

References

1. Howard, R.A., Matheson, J.E.: Influence diagram. In: Howard, R.A., Matheson, J.E. (eds.) The Principles and Applications of Decision Analysis, vol. 2, pp. 721–762. Strategic Decision Group (1981)
2. Kim, J.K., Lee, K.C., Lee, J.K.: Hybrid of neural network and decision knowledge approach to generating influence diagrams. Expert Syst. Appl. **23**, 237–244 (2002)
3. Kim, J.K., Chu, S.C.: Sensitivity analysis in the decision class analysis using neural networks. In: 4th World Congress on Expert Systems, Mexico, pp. 874–879 (1998)
4. Bai, L., Liu, W.Y.: An influence diagram structure learning algorithm based on scoring-search. In: 10th Joint International Computer Conference, pp. 100–104. World Publishing Corporation, Kunming (2004)
5. Ng, A.Y., Russell, S.: Algorithms for inverse reinforcement leaning. In: 17th International Conference on Machine Learning, Stanford, pp. 663–670 (2000)
6. Chajewska, U., Koller, D., Ormoneit, D.: Learning an agent's utility function by observing behavior. In: 18th International Conference on Machine Learning, Williamstown, MA, pp. 35–42 (2001)
7. Nielsen, T.D., Jensen, F.V.: Learning a decision maker's utility function from (possibly) inconsistent behavior. Artif. Intell. **160**(1), 53–78 (2004)
8. Tatman, J.A., Shachter, R.D.: Dynamic programming and influence diagrams. IEEE Trans. Syst. Man Cybern. **20**, 265–279 (1990)
9. Pearl, J.: Probabilistic Reasoning in Intelligence Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers, Los Altos (1988)

A Brief Review of Image Restoration Techniques Based on Generative Adversarial Models

Cai Zhang, Fei Du, and Yungang Zhang(✉)

Department of Computer Science,
Yunnan Normal University, Kunming 650500, Yunnan, China
Jackzhang.cai@foxmail.com, fei.du01@foxmail.com,
yungang.zhang01@gmail.com

Abstract. Images are possibly degraded by various reasons, the typical forms of degradation are: blur, noise, low resolution, and etc. Image restoration techniques try to recover the degraded images to the original images with maximum fidelity. Image restoration is a challenging task and also an import research area in image processing. During the decades, researchers have proposed many restoration methods such as inverse filter, Weiner filter, wavelet analysis, support vector machine, and etc. Recently, deep learning has been increasingly popular among researchers and has obtained remarkable results. In this paper, we briefly review the approaches based on generative adversarial networks (GANs) for image restoration. The typical GANs based restoration methods for image super-resolution, image denoising, image inpainting and image deblurring are introduced and discussed.

Keywords: Image restoration · Degraded image · Deep learning · Neural network · Generative adversarial networks

1 Introduction

The generative adversarial networks (GANs) are proposed by Ian Goodfellow in 2014 [1]. GANs have dramatically sharpened the possibility of AI-generated contents and have drawn active research since they were first described. GANs are neural networks that learn to create synthetic data similar to some known input data. For instance, researchers have generated convincing images from photographs of everything from bedrooms to album covers, and they show a remarkable ability to reflect higher-order semantic logic. GANs consist of two models: a generative model and a discriminative model (Fig. 1).

The generator model takes random input values i.e. dimensional noise vector and transforms them into images through a deconvolutional neural network. On the other hand, the discriminator model is a classifier that determines whether a given image

The project is funded by Natural Science Foundation China No. 61462079 and Yunnan Provincial Education Department Research Grant No. 2018JS143.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 169–175, 2020.
https://doi.org/10.1007/978-981-32-9244-4_24

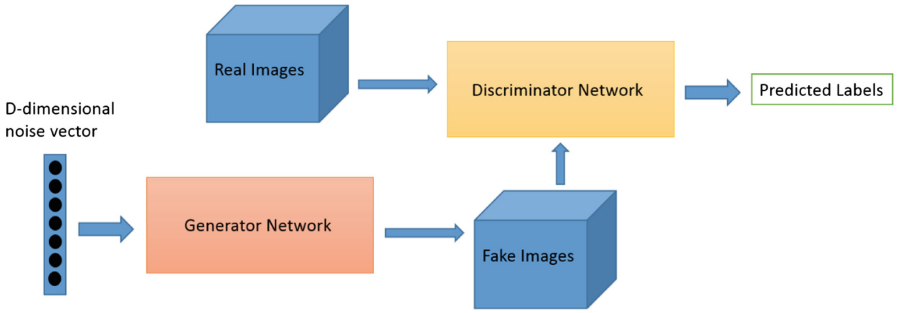


Fig. 1. The GANs architecture

looks like a real world image from dataset or like a generated image. This is basically a binary classifier that will take the form of a normal convolutional neural network (CNN).

A GAN model can be treated as a combination of a counterfeiter (the generator) and a cop (the discriminator) in a game of cat and mouse, where the counterfeiter tries to pass fake notes and the cop tries to detect them out. Both of them are dynamic; i.e. either the generator or the discriminator tries to learn the opposing strategies and escalate themselves. In recent years, researchers have proposed various restoration methods based on GANs model. These models have been applied in many fields of image restoration such as image super-resolution, image inpainting, image deblurring and image denoising.

2 Image Super-Resolution

Image super-resolution is a typical application of GAN in image restoration. Richard Herbert proposed a similar to model of deep belief network: EnhanceGAN [2], the model is a combination of adversarial network and auto-encoder which is used to generate the 8×8 image into the 32×32 image for super resolution. EnhanceGAN has three stages of training, similar to a deep belief network, how even the combined adversarial nets and autoencoders take place of the Restricted Boltzmann Machines. The architecture of EnhanceGAN's stage one and stage two is shown in Fig. 2.

Johnson et al. [3] proposed a perceptual loss function and used in image style transfer and super resolution, the model obtains good results. Ledig et al. [4] proposed a method based on GAN to transform a low-clear image into a HD image with rich details. In [5] the authors use VGG network as the discriminator and parameterized residual network [6] as the generator. The paper adds adversarial loss on the basic of perceptual loss to constrain super-resolution results satisfying natural image distribution principle so that super-resolution image gains very realistic detail effect. However, the disadvantage of this method is that since adversarial loss considers the overall distribution of natural images, regardless of the specific input image (i.e., the test image), the restored image details may not be faithful to the original image, similar to "fabricating" false details.

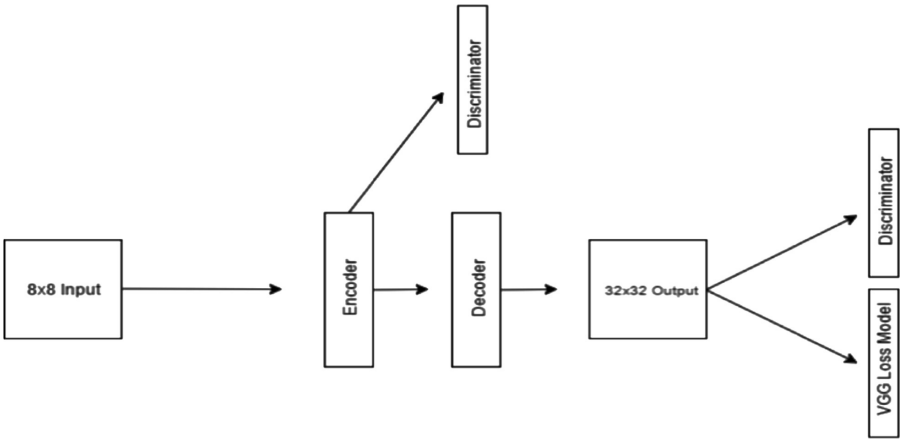


Fig. 2. Diagram of EnhanceGAN's stage one and two models

You et al. [7] proposed a semi-supervised deep learning approach to accurately restore low-resolution (LR) counterparts into high-resolution (HR) computed tomography (CT) images. Especially, with the generative adversarial network (GAN) as the building block, which enforce the cycle-consistency in terms of the Wasserstein distance to establish a nonlinear end-to-end mapping from LR input image to the denoised and deblurred HR outputs, the architectures of the generators and the discriminators are shown in Fig. 3. The model is accurate, efficient and robust for super-resolution image restoration from noisy LR input images. Mustafa et al. [8] proposed a computationally efficient image enhancement approach which provides a strong defense mechanism to effectively mitigate the effect of such adversarial perturbations. The proposed scheme is simple however it has various merits.

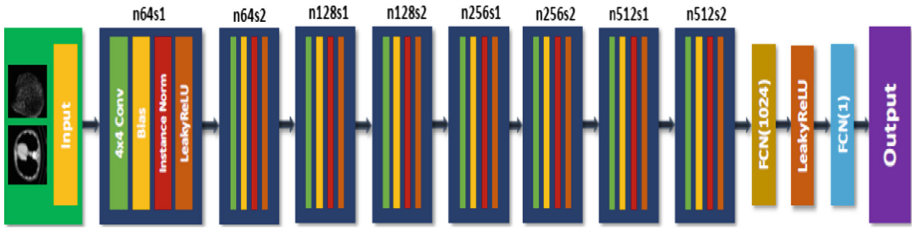


Fig. 3. The architecture of the discriminators, where n stands for the number of convolutional kernels, and s stands for stride. i.e., n32s1 means the convolutional layer of 32 kernels with stride 1 [7].

3 Image Inpainting

In image inpainting field, GAN is widely applied by many researchers. Yeh et al. use deep convolutional neural network for generative adversarial learning, deep convolutional generative adversarial network (DCGAN) [9] uses two loss functions to achieve image restoration where the context loss is used to keep similarity between the degraded image and the restored image and the perceptual loss is used to control visual continuity of the restored image. Pathak et al. [10] proposed CC-GAN, a context constraint keep similarity between the degraded image and the restored image and perceptual loss used to control visual continuity of the restored image. Different from DCGAN, CC-GAN uses semi-supervised learning to finish the recovery of image. Emily Denton et al. [11] use context encoder to get features of image, the decoder is used to accomplish image restoration with learned features, whereas the learning process of the encoder is the pre-training process of the model. Iizuka et al. [12] proposed a model with one generator and two discriminators for image restoration, where the generator is used to fill missing content of image, a discriminator is for distinguishing global authenticity, and the other one is used to distinguish validation padding of local pixels. The latest application of the generated adversarial network for image restoration can be found in the literature Li et al. [13]. The authors use the depth generation network, the local and global loss functions to complete the repair of the missing image of the face.

Vitoria et al. [14] tackled image inpainting problem by using the available visual data and incorporating image semantics through generative models. Their work include mainly two aspects: first, they learn a data latent space by training an improved version of the Wasserstein GAN. Second, a new optimization loss is combined with the learned semantic information to inpaint whose minimization infers the missing content conditioned by the available data.

Lahiri et al. [15] present several architectural and optimization recipes for generative adversarial network (GAN) based facial semantic inpainting. Since current benchmark models are sensitive to the initial solutions of non-convex optimization principle of GAN based inpainting, the authors proposed an end-to-end trainable parametric network to exactly start from good original solutions leading to more photo realistic reconstructions with significant optimization speed up.

Since insufficient cognition understanding, existing learning-based methods often produce artifacts and fallacious textures. Xiao et al. [16] proposed deep inception learning to promote advance feature representation and enhance model learning capacity for local patches. Nazeri et al. [17] developed a novel method for image inpainting that does a better job of reproducing filled regions with fine details. They proposed a two-stage adversarial model EdgeConnect that consists of an edge generator and an image inpainting network.

4 Image Deblurring

In the field of image deblurring, although the use of convolutional neural networks has been extensive, the application of GAN in image deblurring has just begun. Kupyn et al. [18] proposed a model structure: deblurGAN, the authors use the same network structure as in [3]. They use two loss functions for image deblurring which are content loss and adversarial loss. However, the demerit of the method is that the deconvolution operation in the method will bring checkboard artifacts in the resulting images.

Xin et al. [19] proposed a Scale-recurrent Network (SRN-DeblurNet) for deblurring task. This network structure is simpler and has smaller number of parameters compared with many recent learning-based approaches. Yuan et al. [20] proved that the encoder-decoder architecture can provide better results for image deblurring task. Then they propose a novel end-to-end learning model which refines GAN by many novel training strategies so as to solve the problem of deblurring. Bai et al. [21] proposed a graph-based blind image deblurring algorithm by transforming an image patch as a signal on a weighted graph.

5 Image Denoising

Zhang et al. [22] used CGAN's network structure and refined perceptual loss function to remove raindrops from the image. This method outperforms the methods of removing noise using CNN and DCGAN. However, this method has the disadvantage of enhancing the white circular particles in the image so that some raindrops cannot be eliminated; secondly, the method still produces a checkerboard effect in the resulting image. Lefkimmiatis et al. [23] introduced a novel network architecture for learning discriminative image models that can be used to tackle the problem of grayscale and color image denoising.

Ren et al. [24] proposed a DN-ResNet, which is a deep neural network in the form of several residual blocks (ResBlocks). DN-ResNet is more accurate and more computationally efficient than the state-of-the-art denoising networks by using cascade training. An enhanced convolutional neural denoising network (ECNDNet) is proposed by Tian et al. [25] for image denoising, the residual learning and batch normalization (BN) techniques are used to address the problem of training difficulties and accelerate the convergence of the network.

6 Conclusions

Image restoration techniques are used to restore degraded image into the original image as much as possible. Image quality can be dramatically improved by using suitable restoration methods so that improve the use value of images. In this paper, the image restoration methods based on the GAN model are highlighted, which are the most advanced and efficient restoration methods. This is of great value and reference for researchers who study image restoration based on noise and blur model.

References

1. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Bing, X., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: International Conference on Neural Information Processing Systems (2014)
2. Herbert, R.: Faces from noise: super enhancing 8 × 8 images with enhance GAN (2016). https://medium.com/@richardherbert/faces-from-noise-super-enhancing-8x8-images-with-enhancegan-ebda015bb5e0
3. Johnson, J., Alahi, A., Feifei, L.: Perceptual losses for real-time style transfer and super-resolution. In: European Conference on Computer Vision (2016)
4. Ledig, C., Theis, L., Huszar, F., Caballero, J., Aitken, A., Tejani, A., Totz, J., Wang, Z., Shi, W.: Photo-realistic single image super-resolution using a generative adversarial network. In: CVPR (2016)
5. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: ICLR (2015)
6. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: CVPR 2016, pp. 770–778. IEEE (2016)
7. You, C., Li, G., Zhang, Y., Zhang, X., Shan, H., Ju, S., Zhao, Z., Zhang, Z., Cong, W., Vannier, M.W.: CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE) (2018). arXiv:1808.04256v3 [cs.CV]
8. Mustafa, A., Khan, S.H., Hayat, M., Shen, J., Shao, L.: Image super-resolution as a defense against adversarial attacks (2018). arXiv:1901.01677v1 [cs.CV]
9. Yeh, R., Chen, C., Lim, T.Y., Hasegawa-Johnson, M., Do, M.N.: Semantic image inpainting with perceptual and contextual losses (2016). arXiv:1607.07539v2 [cs.CV]
10. Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., Efros, A.A.: Context encoders: feature learning by inpainting. In: IEEE Conference on Computer Vision Pattern Recognition (2016)
11. Denton, E., Gross, S., Fergus, R.: Semi-supervised learning with context-conditional generative adversarial networks (2017). arXiv:1611.06430v1 [cs.CV]
12. Iizuka, S., Simo-Serra, E., Ishikawa, H.: Globally and locally consistent image completion. ACM Trans. Graph. **36**, 1–14 (2017)
13. Li, Y., Liu, S., Yang, J., Yang, M.H.: Generative face completion (2017). arXiv:1704.05838v1 [cs.CV]
14. Vitoria, P., Sintes, J., Ballester, C.: Semantic image inpainting through improved Wasserstein generative adversarial networks (2018). arXiv:1812.01071v1 [cs.CV]
15. Lahiri, A., Jain, A., Nadendla, D., Biswas, P.K.: Improved techniques for GAN based facial inpainting (2018). arXiv:1810.08774v1 [cs.CV]
16. Xiao, Q., Li, G., Chen, Q.: Deep inception generative network for cognitive image inpainting (2018). arXiv:1812.01458v1 [cs.CV]
17. Nazeri, K., Ng, E., Joseph, T., Qureshi, F., Ebrahimi, M.: EdgeConnect: generative image inpainting with adversarial edge learning (2018). arXiv:1901.00212v2 [cs.CV]
18. Kupyn, O., Budzan, V., Mykhailych, M., Mishkin, D., Matas, J.: DeblurGAN: blind motion deblurring using conditional adversarial networks (2017). arXiv:1711.07064v3 [cs.CV]
19. Xin, T., Gao, H., Yi, W., Shen, X., Wang, J., Jia, J.: Scale-recurrent network for deep image deblurring (2018). arXiv:1802.01770v1 [cs.CV]
20. Yuan, Q., Li, J., Zhang, L., Wu, Z., Liu, G.: Blind motion deblurring with cycle generative adversarial networks (2018). arXiv:1901.01641v2 [cs.CV]
21. Bai, Y., Cheung, G., Liu, X., Wen, G.: Graph-based blind image deblurring from a single photograph. IEEE Trans. Image Process., 1 (2018)

22. Zhang, H., Sindagi, V., Patel, V.M.: Image de-raining using a conditional generative adversarial network (2017). arXiv:1701.05957v1 [cs.CV]
23. Lefkimmiatis, S.: Universal denoising networks: a novel CNN-based network architecture for image denoising (2017). arXiv:1711.07807v2 [cs.CV]
24. Ren, H., El-khamy, M., Lee, J.: DN-ResNet: efficient deep residual network for image denoising (2018). arXiv:1810.06766v1 [eess.IV]
25. Tian, C., Xu, Y., Fei, L., Wang, J., Wen, J., Luo, N.: Enhanced CNN for image denoising (2018). https://arxiv.org/abs/1810.11834

A Priority Heuristic Correlation Technique for Decision Tree Pruning

Yu Xiang[1(✉)] and Li Ma[2]

[1] School of Information and Technology,
Yunnan Normal University, Kunming, China
iamlionx@126.com
[2] Library of Kunming, University of Science and Technology, Kunming, China

Abstract. The main purpose of this study is to introduce a priority heuristic correlated information processing technique into decision support tree pruning problem by using behavior predicting and reasoning approach along with the probability correlation analysis framework. To verify its validity, evidence from some important tests is illustrated and discussed, and the results suggest that with this new technique, the size of decision tree can be structurally reduced without losing any predictive accuracy and make it better account for real-life decision-making problems.

Keywords: Decision tree pruning · Priority heuristic · Probability correlation

1 Introduction

The most commonly arisen problem in all kind of decision tree building algorithms is to determine the optimal size of the objective tree [1]. If a tree is too large, it may exit potential risks overfitting the training data by converging too early, and badly generalizing to new data samples. If a tree is too small, it may hardly represent important structural information about the sample space, and lack of enough decision support capabilities. For these algorithms it's very hard to know when should stop growing the tree, because they can barely tell that whether a single additional node will significantly improve the decision accuracy and reduce uncertainty. This is somehow named as the horizon effect. A feasible strategy called backward pruning or post-pruning is to grow the tree fully expanded until its leaves merely contain a small number of events or decision information, then prune some branches and retrospectively reduce the size by remove redundant nodes that do not provide enough decision-making power [2]. Another feasible strategy called forward pruning or pre-pruning is to stop the tree growing excessively by controlling the iteration numbers and convergence rate properly [3].

There are various techniques for decision tree pruning that differ in the measurement used to optimize decision making performance. To sum up, they can be divided into two pruning fashions, the top down which traverses nodes and trim subtrees starting at the root, or the bottom up which starts at the leaves. The most frequently used techniques include reduced error pruning, pessimistic error pruning, minimum error pruning, cost complexity pruning, and critical value pruning, all of these methods

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 176–182, 2020.
https://doi.org/10.1007/978-981-32-9244-4_25

always attempt to systematically maximize or minimize certain integrative decision value of information over selections to provide decision path preference [4]. Distinctively, in this article, we dedicate to propose a new method by introducing heuristic framework for decision tree pruning under risk situations based on the good-enough rule and verify its validity with several practical tests.

2 Priority Heuristic

As one of the most influential non-integrative descriptive model of human risky decision making behavior, the priority heuristic (*PH*) model is used as a variant of the lexicographic semiorder model which compares each path of the decision tree on the first attribute, and if the differences between them are less than a threshold, then go on to examine the differences on the second attribute; if that attribute differs them by less than a threshold as well, go on to the third, and go on, at the end of the procedure *PH* chooses a decision path with the most attractive outcome [5], such models are considered fast and frugal because choices can be made, in some cases, without examining all of the information, but *PH* naturally has its additional features and assumptions [6].

For decision tree examination the threshold of stopping rules can always be assumed to be one tenth of the largest outcome on current attribute. *PH* model is an excellent example of using good enough principle which means choosing a simple solution rather than an elaborate one [7]. The good enough does not mean second best or just right, it means acknowledging that, in an emergency response, adopting a quick and simple approach to impact measurement and accountability may be the only practical possibility [8].

3 Pruning with Priority Heuristic

For a practical decision tree classification instance with well selected deterministic data, any random sample in the training data set can always be classified correctly from its known attributes. However, in numerous real-life decision-making problems there unexpectedly may be a certain degree of uncertainty hidden in the data. This uncertainty empirically derives from two different sources. One is that the value space and the range of an attribute or sample class may be incorrectly measured or limited, and the worst of all they may be missed at the data preprocessing phase. The other one is that although some extraneous factors are not recognized or involved, they may still evidently affect the classification results. When some decision tree constructing algorithms classify such data, the resulting trees tend to be very large and risk overfitting. Unfortunately, large tree size does not equal better decision accuracy in most cases.

So, in this article, we would like to propose both *PH* model and systematic pruning methods to be activated simultaneously in a probability correlation framework that Decision-making attributes are examined by utilizing *PH* processing, while probabilities and correlations are calculated by utilizing systematic processing. Put differently, they are treated as complementary approaches in one decision tree pruning process to observably reduce the size of the tree without losing any decision accuracy, so even to

increase the accuracy in some cases. For a completely constructed decision tree T, the entropy H of a node d in T is given by:

$$H(d) = -\sum_i p_i \log_2(p_i) \tag{1}$$

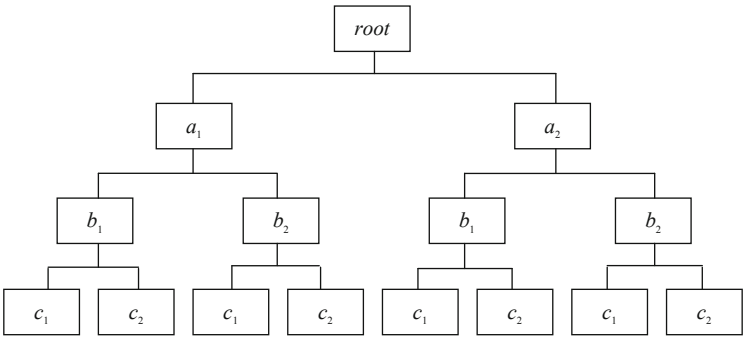


Fig. 1. Binary Decision Tree T

In formula (1) where p_i is the probability of d being examined with the training data samples. Without loss of generality, consider T a binary decision tree (shown in Fig. 1), the probability entropy of all possible decision-making path in T (except all paths start at the root node) are calculated as follows:

$$P_J(a_1, b_1, c_1) = -p(a_1, b_1, c_1) \log_2 p(a_1, b_1, c_1) = p_1 \tag{2}$$

$$P_J(a_1, b_1, c_2) = -p(a_1, b_1, c_2) \log_2 p(a_1, b_1, c_2) = p_2 \tag{3}$$

$$P_J(a_1, b_2, c_1) = -p(a_1, b_2, c_1) \log_2 p(a_1, b_2, c_1) = p_3 \tag{4}$$

$$P_J(a_1, b_2, c_2) = -p(a_1, b_2, c_2) \log_2 p(a_1, b_2, c_2) = p_4 \tag{5}$$

$$P_J(a_2, b_1, c_1) = -p(a_2, b_1, c_1) \log_2 p(a_2, b_1, c_1) = p_5 \tag{6}$$

$$P_J(a_2, b_1, c_2) = -p(a_2, b_1, c_2) \log_2 p(a_2, b_1, c_2) = p_6 \tag{7}$$

$$P_J(a_2, b_2, c_1) = -p(a_2, b_2, c_1) \log_2 p(a_2, b_2, c_1) = p_7 \tag{8}$$

$$P_J(a_2, b_2, c_2) = -p(a_2, b_2, c_2) \log_2 p(a_2, b_2, c_2) = p_8 \tag{9}$$

In Eqs. (2) to (9) where $p(x, y, z)$ is the joint probability of a decision-making path which starts at node x, passes through node y and ends at node z. Furthermore, considering $C_O(x, y)$ as the probability correlation function defined as follows:

$$C_o(a_1, b_1) = p_1 + p_2 - p_3 - p_4 - p_5 - p_6 - p_7 - p_8 \tag{10}$$

$$C_o(a_1, b_2) = -p_1 - p_2 + p_3 + p_4 - p_5 - p_6 - p_7 - p_8 \tag{11}$$

$$C_o(a_2, b_1) = -p_1 - p_2 - p_3 - p_4 + p_5 + p_6 - p_7 - p_8 \tag{12}$$

$$C_o(a_2, b_2) = -p_1 - p_2 - p_3 - p_4 - p_5 - p_6 + p_7 + p_8 \tag{13}$$

$$C_o(a_1, c_1) = p_1 - p_2 + p_3 - p_4 - p_5 - p_6 - p_7 - p_8 \tag{14}$$

$$C_o(a_1, c_2) = -p_1 + p_2 - p_3 + p_4 - p_5 - p_6 - p_7 - p_8 \tag{15}$$

$$C_o(a_2, c_1) = -p_1 - p_2 - p_3 - p_4 + p_5 - p_6 + p_7 - p_8 \tag{16}$$

$$C_o(a_2, c_2) = -p_1 - p_2 - p_3 - p_4 - p_5 + p_6 - p_7 + p_8 \tag{17}$$

$$C_o(b_1, c_1) = p_1 + p_2 - p_3 - p_4 + p_5 - p_6 - p_7 - p_8 \tag{18}$$

$$C_o(b_1, c_2) = -p_1 + p_2 - p_3 - p_4 - p_5 + p_6 - p_7 - p_8 \tag{19}$$

$$C_o(b_2, c_1) = -p_1 - p_2 + p_3 - p_4 - p_5 - p_6 + p_7 - p_8 \tag{20}$$

$$C_o(b_2, c_2) = -p_1 - p_2 - p_3 + p_4 - p_5 - p_6 - p_7 + p_8 \tag{21}$$

In more general terms, referencing formulae (10) to (21), if nodes $x, y \in \{a_{i1}, b_{i2}, c_{i3}\}$ and $x, y \notin \{a_{j1}, b_{j2}, c_{j3}\}$, then $C_o(x, y)$ can be defined as:

$$C_o(x, y) = \sum P_J(a_{i1}, b_{i2}, c_{i3}) - \sum P_J(a_{j1}, b_{j2}, c_{j3}) \tag{22}$$

Formula (22) is a sum of all probability entropy which supports the correlation between nodes x and y subtracts a sum which is against them.

$$C_o^{\max} = \frac{1}{10} \max_{i,j=1}^{2} \left(C_o(a_i, b_j), C_o(a_i, c_j), C_o(b_i, c_j)\right) \tag{23}$$

To introduce the main idea of the *PH* model into probability correlation framework, correspondingly, define $C_o^{\max}$ in formula (23) as the threshold which is one tenth of the most significant correlation between any two random nodes (exclude the root node) in the decision tree T, and then we propose the pruning rules for decision tree T as: If the correlation of node pair x and y is more significant than the sum of all other possible correlations of node pairs on the branches that contain at least one of these two nodes, and the correlation is also more significant than $C_o^{\max}$, then trim all intermediate nodes or sections between x and y to make them directly connected. If x and y are directly connected and their correlation is more significant than the sum of all other possible correlations of node pairs on the branches which contain at least one of them, while their correlation is more significant than $C_o^{\max}$, then remove all child branches or subtrees from them.

For example, considering a node pair a_1 and c_1 in decision tree T (as shown in Fig. 1), if we know that $C_o(a_1, c_1) > C_o(a_1, b_1) + C_o(b_1, c_1), C_o(a_1, c_1) > C_o^{\max}$, and $C_o(a_1, b_2) > C_o(a_1, c_2) + C_o(b_2, c_2), C_o(a_1, b_2) > C_o^{\max}$, then T can be pruned to T_{co} (as shown in Fig. 2).

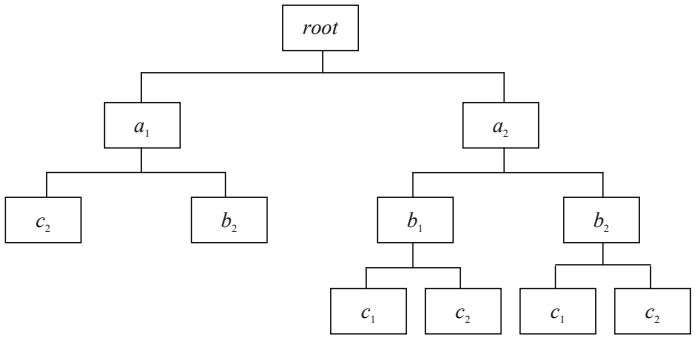


Fig. 2. Pruned Binary Decision Tree T_{CO}

4 The Verification Tests

To verify the validity of this probability correlation pruning framework, a decision tree T is acquired by using *CART* algorithm on the basis of the coronary heart disease data set (known as *CHDD* which is usually used for machine learning, and is also a subset of *CHDAGE.DAT* from Hosmer & Lemeshow, Applied Logistic Regression, copyrighted by John Wiley & Sons Inc.) that represents observations on 9 decision-making attributes and provide 3317 structurally complete positive data samples. In *CHDD*, *train* is the training data sample set and *evalu* is the evaluation data sample set. By contrast, T is pruned to T_{co} by using *PH* probability correlation technique, and both are iterated 50 times to reach their acceptable convergence state.

As the depth of trees are 3, according to Figs. 3 and 4, T_{co} earns a better average decision-making score and will be convergent a little bit earlier than T, and both resulting trees are not overfitted to data samples after stopping the iterating process and growing which makes them convergent without a large number of iterations. As the depth of trees are 4, according to Figs. 5 and 6, T_{co} still has a better average decision-making score and will also be convergent earlier than T, and both resulting trees, as we expected, are not evidently overfitted after stopping the iterating process and growing, but both require more iterations to reach convergence. As the depth of trees are 5, according to Figs. 7 and 8, T_{co} gains a better average decision-making score than T which is observably overfitted at the early phase of the iterating process while T_{co}, in a manner, postpones it. This can be also considered that T_{co} requires less iterations to reach its convergence state which theoretically due to a more optimized structure reduction of the decision tree.

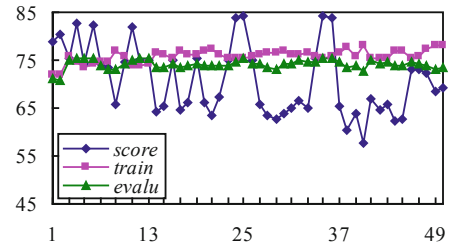


Fig. 3. Test results of T with depth 3

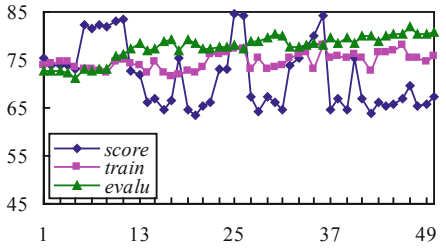


Fig. 4. Test results of T_{CO} with depth 3

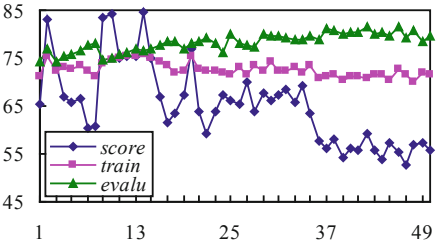


Fig. 5. Test results of T with depth 4

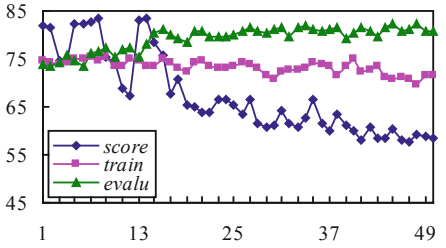


Fig. 6. Test results of T_{CO} with depth 4

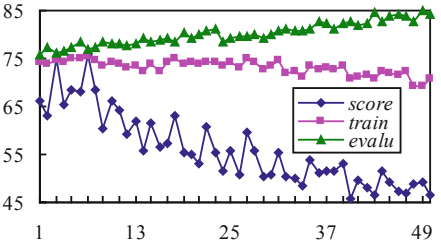


Fig. 7. Test results of T with depth 5

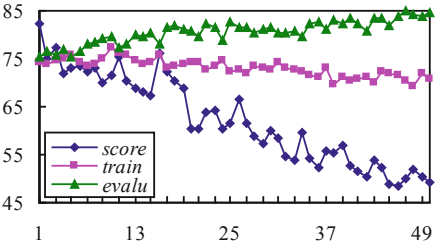


Fig. 8. Test results of T_{CO} with depth 5

5 Conclusion

There are many decision tree pruning methods, some work on constructing more accurate decision trees, and some focus on building decision trees with smaller sizes. While this study aims to propose a probability correlation pruning framework by using both the *PH* model and the frequently-used systematic pruning method as complementary techniques in one decision tree pruning implementation. Moreover, to verify the validity, feasibility and practicability of this new technique, some important test results are demonstrated and discussed which draw a positive conclusion that the *PH* correlation pruning method corresponds with the decision-making strategy for real-life classification problems with more accuracy and smaller tree size.

What deservers more investigation is that the CHDD samples used to run the tests may be more theoretically and empirically correlative than data samples from other fields owing to pathology and ecsomatics reasons. Thus, new practical data samples are required to be tested with this new technique to further verify its validity.

References

1. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Leaning: Data Mining, Inference, and Prediction, pp. 409–432. Springer, New York (2009)
2. Kaminski, B., Jakubczyk, M., Szufel, P.: A framework for sensitivity analysis of decision tree. CEJOR **26**, 135–159 (2018)
3. Shenoy, P.: A new pruning method for solving decision trees and game trees. In: Besnard, P., Hanks, S. (eds.) Uncertainty in Artificial Intelligence, vol. 11, pp. 482–490 (1995)
4. Esposito, F., Malerba, D., Semeraro, G.: A comparative analysis of methods for pruning decision trees. IEEE Trans. Pattern Anal. Mach. Intell. **19**, 476–491 (1997)
5. Eduard, B., Gerd, G., Ralph, H.: The priority heuristic: making choice without trade-offs. Psychol. Rev. **113**, 409–432 (2006)
6. Birnbaum, M.: New paradoxes of risky decision making. Psychol. Rev. **115**, 463–501 (2008)
7. Serlin, R., Lapsley, D.: Rationality in psychological research: the good-enough principle. Am. Psychol. **40**, 73–83 (1985)
8. Passanisi, W., ECB Project: The Good Enough Guide: Impact Measurement and Accountability in Emergencies, pp. 5–6. Oxfam GB, Oxford (2007)

A Rapid Response Approach Applying Edge Computing for Distributed Warehouses in WSN

Yuechun Wang[1], Ka Lok Man[2,3(✉)], Danny Hughes[4], Steven Guan[1], and Prudence Wong[5]

[1] Department of Computer Science and Software Engineering (CSSE), Xi'an Jiaotong-Liverpool University (XJTLU), Suzhou, China
{yuechun.wang, steven.guan}@xjtlu.edu.cn
[2] AI University Research Centre (AI-URC), Department of Computer Science and Software Engineering (CSSE), Xi'an Jiaotong-Liverpool University (XJTLU), Suzhou, China
ka.man@xjtlu.edu.cn
[3] Swinburne University of Technology Sarawak, Kuching, Malaysia
[4] imec-DistriNet, KU Leuven, 3001 Leuven, Belgium
danny.hughes@cs.kuleuven.be
[5] Department of Computer Science, University of Liverpool, Liverpool, UK
p.wong@liverpool.ac.uk

Abstract. This paper presents a rapid response system architecture for the distributed management of warehouses in logistics by applying the concept of edge computing. A tiered edge node architecture is proposed for the system to process computing tasks of different complexity, and a corresponding rapid response algorithm is introduced. A software-defined simulation is done to evaluate the system performances on rapidness and correctness, from which it can be concluded that all pre-defined emergency cases can be detected and responded to within a relatively short period of time.

Keywords: Edge computing · Rapid response algorithm · Wireless Sensor Networks (WSN) · Distributed warehouses management · Logistics

1 Introduction

As a mature Internet of Things (IoT) scenario, Wireless Sensor Networks (WSN) have rapidly proliferated over the last decade. These diversified WSN applications are emerging rapidly, while the quantity of various nodes and platforms in the WSN is increasing exponentially. Wearable human sensor network, the smart home, intelligent logistics and transportation [1], as well as the smart city [2] are typical application scenarios for current WSNs.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 183–189, 2020.
https://doi.org/10.1007/978-981-32-9244-4_26

From the perspective of logistics, the warehouse is a critical scenario for contemporary intelligent logistics applications. For warehouse management, sensed data may be used for two general purposes: one is for cargo management, which includes goods identification (using RFID) and goods tracking (location and movement); the other is for safety management, which refers to the environmental monitoring and data security. The logistics companies who aim to conduct business nationwide need to consider both the centralized global control of the profession on Cloud as well as the management of distributed networks of warehouses locally. At this scale, the traditional WSN plus Cloud mode may lead to either high bandwidth use or latency in undertaking emergency interventions. In short, many applications require both WSN localization and Cloud globalization which cannot be satisfied by a simple WSN-Cloud architecture. In such case, a well-designed WSN-Edge-Cloud system architecture that integrates edge computing features with the WSN-Cloud architecture would solve these problems and improve the efficiency of the business [3, 4].

Being regarded as a relay between the data centre on the Cloud and sensor nodes in the WSN, edge computing nodes extend the Cloud Computing paradigm to the edge of the network in a bidirectional way. On the node-to-cloud direction, edge nodes revolve around local functionality for geographically closer sensing area with the feature of data pre-processing and rapid reaction [5]. These outcomes will be sent to the Cloud selectively, according to the explicit application requirements. In the cloud-to-node direction, edge nodes achieve distributed deployment of the broad class of applications under the macro control of the Cloud and perform the tasks allocated by the Cloud [6].

2 Architecture and Methodology

2.1 Edge Computing-Based Graded System Architecture

The edge computing-based graded system architecture, which can be differentiated from traditional WSN system architecture, consists of three general layers at the vertical direction, which from top to bottom are the Cloud, Edge, and WSN infrastructure as shown in Fig. 1. As a widely accepted environmental sensing infrastructure, sensor nodes in the WSN collect sensing data and track changes of the environment continuously. For better identification and management, sensor nodes in the WSN are logically separated into different areas.

The Edge computing layer is introduced into the system by considering it as the implementation of cloud computing close to the physical environment. The functionality of edge computing is refined into three grades of edge nodes. Grade one and two edge nodes are focused on the data formatting, preliminary data processing for WSN data collection, as well as the execution of tasks and control commands allocated by the

upper layer (higher grade edge nodes or the cloud). Grade three edge nodes contribute to more complex data analysis, which involves data that is potentially useful for prediction and control, as well as generating or relaying control commands from the upper layer to the lower layer.

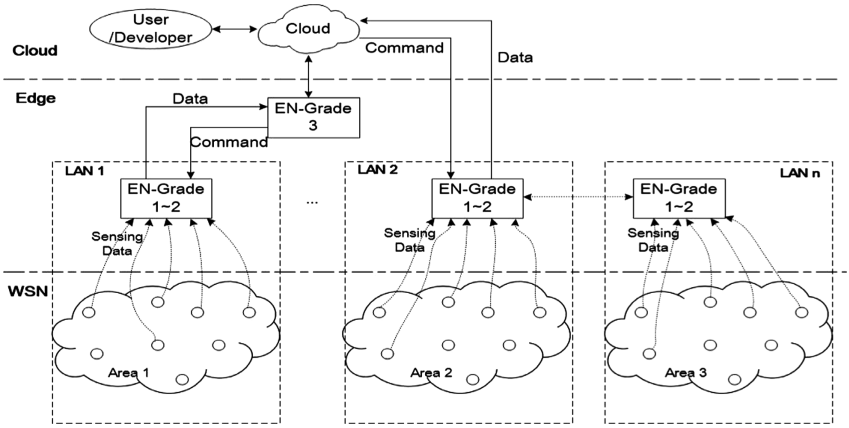


Fig. 1. Edge computing-based tiered system architecture. 'EN' indicates edge node. A smaller EN-Grade number indicates the physically closer location from the edge node to the sensor nodes and sensing devices.

The cloud layer contributes to the centralized analysis of global data and management of the entire network. In addition, the connection between users and the system via the cloud realizes the remote operation and control all areas covered by the terminal devices. For application developers, the system can be accessed via the cloud or edge node for application deployment regarding the deployment requirements and the network condition.

2.2 Rapid Response Algorithm

Within a target monitoring area, there are two primary cases in which sensor nodes may generate abnormal sensing data: one is the sudden environmental change, the other is the error data caused by sensor broken or irruption. A rapid response is only expected to be triggered by the first case, which could save time for emergency interventions and reduce the potential for business losses. In contrast, a rapid response caused by the second case will lead to a waste of resources.

Algorithm 1. Rapid Response Algorithm

Inputs:

Bound threshold: TH_b
Observe period 1: T_1
Data set: $D(n)=\{d(n,i), n\in N, i:\text{timestamp}\}$
Data status: $C(d)=\{\text{true, when } d\geq TH_b; \text{false, when } d<TH_b, d\in D\}$
Node set: $N=\{n_1, n_2, \ldots, n_k, k:\text{number of nodes in the area}\}$
Neighbor node set: $N_{nb}(n)$
Critical percentage: p
Observe period 2: T_c
Data gradient set: $Grad(n)=\{grad(n,i), n\in N, i:\text{timestamp}\}$
Data gradient: $grad(n,i)=[d(n,i+T_c)-d(n,i)]/T_c, n\in N,$ i:timestamp

Process on EN-1

```
if d(n,i)>TH_b:
   observe C(d), where d∈D_nb, D_nb={d(m,j), m∈N_nb(n), j∈(i,i+T_1)}
   if C(d)=true:
      activate Alarm-1
   else :
      generate M=(n,i)
      activate EN-2
```

Process on EN-2

```
observe M
compute Grad(m), where m∈N_nb(n)
if #{grad(m,i)>0}/#{N_nb}>p:
    activate Alarm-2
else :
    if grad(n,i>0):
        activate Alarm-3
    else :
        record data-error
```

The rapid response algorithm proposed in this paper as shown in Algorithm 1 is under the premise of ensuring accuracy, which classifies the urgent cases into three types: Rapid Growth, Slow Growth-Diffusion, and Slow Growth-Non-Diffusion.

Threshold setting is one of the most popular approaches to distinguish abnormal data and normal data. At the Grade-1 edge nodes, the bound threshold (TH_b) is set for real-time comparison. Any sensing data collected by a sensor node that upload to the Grade-1 edge node will be compared with TH_b. There are two cases that may happen at the Grade-1 edge node by comparing real-time sensing data with TH_b: (a) more than one sensor nodes are distinguished as abnormal within a short period (say T_1), (b) abnormal data appear on a single node. For case (a), we consider a Rapid Growth case happened and generate Alarm-1 directly; while for case (b), Grade-1 edge node will trigger Grade-2 edge node with a new generated message which includes abnormal sensor node ID for further computing and judgement. Once Grade-2 edge node be triggered by this message, it starts to analyze the trend of sensing data of both the abnormal node and its neighbor nodes. The trend is measured by computing the

gradient of data in adjacent time point. The percentage of neighbor nodes who have the same trend as abnormal one will decide the urgent case type: if there are more than p (a given percentage) neighbor nodes have the same trend, we consider a Slow Growth-Diffusion case happened and generate Alarm-2. Otherwise, the trend of the abnormal node in the coming period of time (say T_c) decides the urgent case type. A continuous change of sensing data will denote a Slow Growth-Non-Diffusion case happened and an Alarm-3 will be generated; while the stable sensing data indicates that an error has occurred and there will be no alarm message generated.

3 Implementation

In our simulations, there are four cases considered as listed in Table 1. Case 1–3 are corresponding to the three urgent cases types as introduced in Sect. 2.2, which are Rapid Growth, Slow Growth-Diffusion, and Slow Growth-Non-Diffusion. Case 4 indicates unexpected error data occurs on a single node. Each of the cases corresponds to an alarm type, which is generated by the edge node to distinguish the cases.

Table 1. List of all cases for experiments

Cases	Specification	Alarm type
Case 1	Abnormal of environment is observed by a group of nodes (Rapid Growth)	Alarm-1
Case 2	Abnormal of environment is observed by single node, the abnormal is diffusion (Slow Growth-Diffusion)	Alarm-2
Case 3	Abnormal of environment is observed by single node, the abnormal is non-diffusion (Slow Growth-Non-Diffusion)	Alarm-3
Case 4	Error data on single node	None

Corresponding to three real-world scenarios that produced temperature sensing data, which are open flame spread, high-temperature steam leakage and diffusion, as well as node device over temperature, we generate three sets of data by software for the experiments on case 1–3. The error data for case 4 is inserted into a data set simulated the indoor environmental change manually. All the simulated sensing data are sampled once per second during the experiments.

The performances on correctness and rapidness corresponding to sub-figures (a)–(d) in Fig. 2 are tested under four simulated environments. For each test, there are four nodes updated sensing data over 60 s to the edge node simultaneously. Referring to the input parameters listed in Algorithm 1, the 'bound threshold (TH_b)' is set to be 40 while the 'observe period one (T_1)' is set to be 1. The 'observe period two (T_c)' is 5 and the 'critical percentage (p)' is 0.5 in our experiments.

The system response time is pointed out by an arrow on each sub-figure, which corresponding to the timestamp 23 s, 24 s, 25 s, and 13 s. Comparing with the bound threshold line (labelled as *BoundTh* in the figure), it can be observed that all the emergency cases are detected and responded within an 'observe period two'.

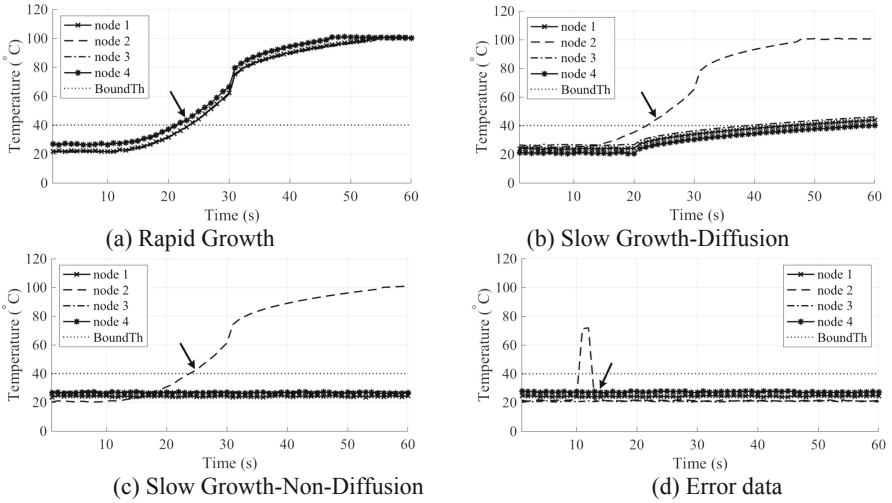


Fig. 2. Simulation on correctness and rapidness performance of the algorithm

4 Conclusion

A rapid response system architecture is proposed in this paper, which involves the concept of edge computing in WSN. From the perspective of distributed warehouse management in logistics, an algorithm for distinguishing and rapidly responding to emergency cases is introduced. Tested by a software-defined simulation, the performance on the correctness and rapidness of the Grade-1 and Grade-2 edge nodes in the system applying the rapid response algorithm shows that all pre-defined emergency cases can be detected and responded within a relatively short period of time.

To implement the entire system architecture as proposed in this paper, a clear direction for future research is the implementation of Grade-3 edge nodes, which potentially focuses on the short-time prediction. Besides the edge computing layer, the interaction and interoperation between the edge and the Cloud is also a valuable direction to extend our research.

Acknowledgement. This project is supported by the Research Development Fund (#RDF14-03-12) and XJTLU Key Programme Special Fund (#KSF-P-02) of the Xi'an Jiaotong-Liverpool University, Suzhou, China.

References

1. Xiao, X., et al.: Developing an intelligent traceability system for aquatic products in cold chain logistics integrated WSN with SPC. J. Food Process. Preserv. **40**(6), 1448–1458 (2016)
2. Mohamed, N., et al.: SmartCityWare: a service-oriented middleware for cloud and fog enabled smart city services. IEEE Access **5**, 17576–17588 (2017)

3. Bonomi, F., et al.: Fog computing and its role in the Internet of Things. In: Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing. ACM (2012)
4. Mihai, V., et al.: Wireless sensor network architecture based on fog computing. In: 5th International Conference on Control, Decision and Information Technologies (CoDIT). IEEE (2018)
5. Zhang, G., Li, R.: Fog computing architecture-based data acquisition for WSN applications. China Commun. **14**(11), 69–81 (2017)
6. Carrega, A., et al.: A middleware for mobile edge computing. IEEE Cloud Comput. **4**(4), 26–37 (2017)

SwarMotion: A 3D Point Cloud Video Recording Tool for Classification Purposes

Diego Monteiro[1], Jialin Wang[1], Hai-Ning Liang[1(✉)], Nilufar Baghaei[2], and Andrew Abel[1]

[1] Department of Computer Science and Software Engineering, Xi'an Jiaotong-Liverpool University, Suzhou, China
haining.liang@xjtlu.edu.cn
[2] Department of Information Technology, Otago Polytechnic Auckland International Campus, Auckland, New Zealand

Abstract. Modern life is ever more reliant on computers being able to classify the world around them and computer vision is one of the ways computers do it. Nowadays, due to the advent of reliable and low-cost range sensors like Kinect which provide useful 3D data to feed prediction systems with a new dimension of useful information, computer vision is taking a new step with research demonstrating the potential that this kind of data has. However, very little research has been done using spatiotemporal Point Cloud data (PC-Videos). One reason might be the lack of datasets containing PC-Videos. In this paper, we propose SwarMotion, a multimodal recording tool focused on the acquisition of PC-Videos.

Keywords: Point Cloud · Video · Spatiotemporal data · Classification

1 Introduction

Modern life is ever more reliant on computers being able to classify the world around them. We trust computers to give us suggestions about the media they categorize as desirable to our individual tastes. We rely on machines to identify fraudulent transactions in our credit cards. We hope virtual avatars will express emotion when talking to us [1]. We talk to our phones expecting them to classify air vibrations as words and bring back some sort of relevant information. We even use computer to try to understand our minds [2, 3]. And one of our greatest expectations for computer classification is computer vision [4].

Computer vision has been around for a while now, using mainly bi-dimensional (2D) images for object classification [5]. Even though, classification using bi-dimensional information has been shown to be particularly successful in certain problem sets, using 2D images has proved difficult in some situations such as cluttered backgrounds or even environments with light variations [6].

Recently, due to the advent of reliable and low-cost range sensors (e.g. RGB-D cameras such as Microsoft Kinect and Intel RealSense), which provide useful 3D data to feed prediction systems with a new dimension of useful information, there has been a spike in the number of research papers studying object classification using 3D Point

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 190–195, 2020.
https://doi.org/10.1007/978-981-32-9244-4_27

Cloud (PC) data [7]. These studies so far have demonstrated the potential that this kind of data has, especially in real word applications [6–11].

Nevertheless, computer vision that analyzes spatiotemporal data is still mainly focused on 2D frames (i.e. videos) [6]. Very few works use spatiotemporal PC (PC-Videos), one reason might be the lack of datasets containing PC-Videos [12]. Thus, to create datasets of PC-Videos, it is ideal to have a tool to help the creation of such datasets, some sort of PC-video capture tool and editor that are integrated with the capture of other temporal data. In this paper, we propose such a tool that we call SwarMotion.

SwarMotion is a multimodal recording tool focused on the acquisition of PC-Videos. In this paper we will explain similar technologies and their limitations, explain the importance of SwarMotion, its development and current state, and the reason to have this recording tool be multi-modal rather than focusing on PC alone.

2 Related Work

To determine the necessary characteristics of the tool we analyzed the recent works involving computer vision and other machine learning techniques that could benefit from PC databases.

We observed a tendency to use some form of 3D-CNN (3D-Convolutional Neural Networks) to analyze PC and other tri-dimensional data, for example the works of [13] and others on NormalNet [4], PointCNN [11] and PointNet [14]—this latter work is not to be confused with another body of papers also dubbed PointNet that propose a non-CNN method [8, 9]. Another, group of papers uses depth in a similar way to color in a 2D frame; they also use a CNN for their analyzes [15].

We did observe a few papers studying movement in tridimensional data [5, 6], which also use CNN; one of them using only PC to recognize human gestures. Further, we observed that, to improve their success rate, other CNN applications use multimodal data (i.e. non-visual data) for their classification [7, 16]. Other machine-learning techniques using multimodal speech recognition for example present better results than audio or visual alone [17, 18].

Based on this review, we can see that the dataset creation tool must be able to synchronously edit all incoming data. It should be able to save the PC data as well as the accompanying data at the same frame rate. Raw data should be available for input and easier editing, or at least some sort of lossless data compression format.

In order to determine the need for a new tool that is aimed at data collection for PC machine learning classification purposes and based on the previously observed papers, we have established a set of baseline features that should be available in the existing tools. We started then the investigation using the following guidelines:

- The tool should be able to record color 2D data along the PC data (i.e. multimodal data);
- The tool should be able to record infrared 2D data along the PC data (i.e. multi-modal data);
- The tool should be able to record Audio along the PC data (i.e. multimodal data);

- The tool should be able to create temporal regions of interest (i.e. edit length of the PC-Video);
- The tool should be able to reproduce collected data to check recorded data;
- The tool should be able to leave the raw data accessible for other kinds of future modification (i.e. saved data not only in proprietary formats); and
- The tool should be able to connect to low cost range sensors (i.e. Microsoft Kinect, Intel RealSense).

2.1 Existing Software

The starting point of the investigation was works that studied Point Cloud and looked for mentions of how data was captured. The majority used online available data sets. One used their own system to capture a sequence of point clouds. Others were not compatible with low cost range sensors and were adapted only for the state-of-the-art technology.

We then searched for commercial solutions because we believed those would have a greater chance of presenting all the requirements, specially tools from Microsoft, since the company has its own range sensor (i.e. Microsoft Kinect). The keywords used for the search were, "Point Cloud Recorder", "Point Cloud Video", "Point Cloud Editor", "Point Cloud Multimodal Recorder", "Volumetric Recording", "Volumetric Film".

We encountered a variety of PC editors including tools to create PC from pictures using photogrammetry, and programs to import laser scanned data as PC. Mostly these found programs were used to create 3D models from the Point Clouds.

Later we looked at opensource solutions that could be available from other researchers who want to create PC video datasets. The search criteria were the same. However, the criteria for exclusion were the following, tools that did not present any sort of documentation and tools that have been discontinued (i.e. did not present an update in over 5 years). On Table 1 we can see a sample of the encountered software and their limitations regarding the requirements.

Table 1. Existing software evaluation following the stipulated criteria

Software name	Company/Developer	Limitations
Web Point Cloud Viewer	AMC Bridge	Static Point Cloud
Cloud Chamber	Spar3D	Static Point Cloud
PointCloudUnity	Miroslav Byrtus*	Static Point Cloud
PointClouds	Edgar077*	Static Point Cloud
RealityCapture	Capturing Reality	Static Point Cloud
KiwiViewer	Kitware	Visualization Tools
Voxxlr	Geist Software Labs	Storage Tool
TrueDepth	Apple	Device Bound
EF EVE	EF EVE	Proprietary Format
Brekel Pointcloud	Brekel	No IR Capture
Depthkit	Depthkit.tv	No Audio Capture, No raw data

Even though we did find several software solutions for the treatment of PC, none fulfilled all the stipulated requirements. The one that reached the closest to the ideal tool we proposed was Brekel PointCloud, but it still lacks the possibility of simultaneously record other temporal data. Thus, we believe the creation of a new tool would be adequate for the advancement of classification techniques.

3 Development and Results

Based on the requirements we proposed the data flow presented in Fig. 1. The sensors connect to the machine and simultaneously send the data stream. The data stream is processed by SwarMotion that generates the PC and make it available for the user to create an STBox (Spatiotemporal bounding box [19]) on the ROI (Region of Interest). The STBox is then converted into the pertinent formats and stored.

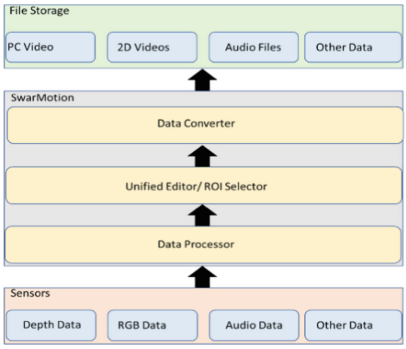


Fig. 1. SwarMotion dataflow concept

SwarMotion was developed using C# on the Unity Engine to facilitate the visualization process. Because of the multimodal requirements, the user can record 2D data and audio to be played along the PC. The initial version of SwarMotion (see Fig. 2) allows the user to record RGB-D data, infrared and sound data and to select the temporal region of interest. The files are currently saved in PLY, PNG, and WAV. This initial version is only designed to use the Kinect and due to its limitations for audio recording a separate microphone must be used.

The recorder, player and editor share the same GUI. The player and editor part can be easily deployed in all platforms which are supported by Unity3D. Such as Android, IOS, video game console and WebGL. This means the recorded files can be easily spread like normal video. And it has a favorable expansibility due to the Unity3D.

Fig. 2. A screenshot of the SwarMotion recording and video editing screen

4 Summary and Future Work

In this paper we explained one direction that machine learning is taking and showed that the current tools and datasets are not enough for the continue development of the field. We then proposed and developed a tool for the creation of new Point Cloud (PC) datasets that consider the temporal and multimodal aspects of the data.

As a future work we plan to develop the tool further allowing for the spatial editing of the data as proposed by us, the connectivity with different RGB-D devices, the recording of other kinds of data, such as Ultra-Violet and temperature, which might help increase the precision of the machine's predictions.

Acknowledgements. This research was partially funded by the Xi'an Jiaotong-Liverpool University (XJTLU) AI University Research Centre (AI-URC), XJTLU Key Program Special Fund (#KSF-A-03 and #KSF-02) and the XJTLU Research Development Fund.

References

1. Monteiro, D., Liang, H.-N., Wang, J., Wang, L., Wang, X., Yue, Y.: Evaluating the effects of a cartoon-like character with emotions on users' behaviour within virtual reality environments. In: IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR), pp. 229–236. IEEE (2018)
2. Monteiro, D., Liang, H., Zhao, Y., Abel, A.: Comparing event related arousal-valence and focus among different viewing perspectives in VR gaming. In: Proceedings of the 9th International Conference on Advances in Brain Inspired Cognitive Systems, BICS 2018, pp. 1–10. Springer (2018)
3. Monteiro, D.V., Liang, H., Baghaei, N.: Correlating gamers' brainwaves to their subjective feelings in virtual reality games under different viewing perspectives. In: Proceedings of 16th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry (VRCAI 2018), Hachioji (2018)

4. Wang, C., Cheng, M., Sohel, F., Bennamoun, M., Li, J.: NormalNet: a voxel-based CNN for 3D object classification and retrieval. Neurocomputing **323**, 139–147 (2019)
5. Owoyemi, J., Hashimoto, K.: Spatiotemporal learning of dynamic gestures from 3D Point cloud data (2018)
6. Ji, S., Xu, W., Yang, M., Yu, K.: 3D Convolutional neural networks for human action recognition. IEEE Trans. Pattern Anal. Mach. Intell. **35**, 221–231 (2013)
7. You, H., Feng, Y., Ji, R., Gao, Y.: PVNet: a joint convolutional network of point cloud and multi-view for 3D shape recognition (2018)
8. Qi, C.R., Yi, L., Su, H., Guibas, L.J.: PointNet++: deep hierarchical feature learning on point sets in a metric space (2017)
9. Qi, C.R., Su, H., Mo, K., Guibas, L.J.: PointNet: deep learning on point sets for 3D classification and segmentation. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017)
10. Garcia-Garcia, A., Gomez-Donoso, F., Garcia-Rodriguez, J., Orts-Escolano, S., Cazorla, M., Azorin-Lopez, J.: PointNet: a 3D convolutional neural network for real-time object class recognition. In: Proceedings of the International Joint Conference on Neural Networks, October 2016, pp. 1578–1584 (2016)
11. Li, Y., Bu, R., Sun, M., Chen, B., Wu, W., Di, X., Chen, B.: PointCNN. Adv. Neural. Inf. Process. Syst. **31**, 828–838 (2018)
12. Fu, H., Xu, D., Lin, S.: Object-based multiple foreground segmentation in RGBD video. IEEE Trans. Image Process. **26**, 1418–1427 (2017)
13. Huang, J., You, S.: Point cloud labeling using 3D convolutional neural network. In: 23rd International Conference on Pattern Recognition (ICPR), pp. 2670–2675 (2016)
14. Sein-Echaluce, M.L., Fidalgo-Blanco, Á., Alves, G.: Technology behaviors in education innovation. Comput. Hum. Behav. **72**, 596–598 (2017)
15. Schwarz, M., Schulz, H., Behnke, S.: RGB-D object recognition and pose estimation based on pre-trained convolutional neural network features. In: IEEE International Conference on Robotics and Automation (ICRA), pp. 1329–1335. IEEE (2015)
16. Feng, W., Guan, N., Li, Y., Zhang, X., Luo, Z.: Audio visual speech recognition with multimodal recurrent neural networks. In: International Joint Conference on Neural Networks (IJCNN), pp. 681–688 (2017)
17. Huang, J., Potamianos, G., Neti, C.: Improving audio-visual speech recognition with an infrared headset, pp. 175–178 (2003)
18. Yoshinaga, T., Tamura, S., Iwano, K., Furui, S.: Audio-visual speech recognition using lip movement extracted from side-face images. In: Proceedings of Audio-Visual Speech Processing, pp. 117–120 (2003)
19. Monteiro, D.V., Ferreira, K.R., Santos, R.: An algorithm to discover partners in trajectories (2017)

Mobile Robot-Based Measuring and Mapping of Gas Distribution in Surface Soil

Zhubing Lei, Muchun Zhou, Jian Wang(✉), and Baochuan Fu

School of Electronic and Information Engineering, Suzhou University of Science and Technology, Suzhou 215009, China
{zhblei,mczhou,wangjiansuzhou,fubc}@usts.edu.cn

Abstract. Laws and policies require that landfills and disposal sites be monitored to protect public health and safety from potential explosion hazards owing to underground gas migration and diffusion. This paper aims to measure and map gas distribution in surface soil. Since some current algorithms have their own advantages and disadvantages, several methods are integrated into a compound algorithm to measure gas concentration, to track odour and localize gas source. Simulations are performed to validate the availability of the proposed method. To be aware of whole gas distribution field, a mobile robot governed by the proposed method measured the concentration values step by step and generated a sequence of small concentration maps. Then by revolving the small maps the whole gas concentration map is built. Results demonstrated that the mapped gas concentration field is satisfied.

Keywords: Gas distribution · Mapping · Measuring · Mobile robot · Tracking

1 Introduction

Mapping gas concentration field in surface soil of ground has applications in many realms such as exploring landfill site, monitoring underground pipelines and storage tanks, and searching for chemical substance or pollutants. It contains two main procedures: the first is tracking an odor/gas and measuring its concentration, the second is mapping the concentration field. Mobile robot with sensors traveling on ground is usually used to detect the concentrations of gas in soil, to track the odor, and to localize the source. Abaci et al. [1] began the research and application in 1992 and Russell suggested the Hexagon-path Algorithm [2, 3] in 2003 and 2004. Since then many products [4–9] have improved the classic algorithm in some features.

For a practical task of measuring and mapping gas concentration in surface soil, there are two problems needing being solved satisfactorily. One is an algorithm which can fit for whole process of search. Since gas concentration distribution appears different patterns in different regions, it is not easy to use just one simple algorithm for whole process. The existing methods, which are effective in detecting and tracking gas, are optimal just in a certain part of the whole process. The other is mapping gas concentration field by mobile robot. There were many studies on mapping mercury's distribution in soil using data sampled by hand [10] and off-line work of mapping pollutants' distribution in soil using on-line measured data [11]. The work on creating a

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 196–206, 2020.
https://doi.org/10.1007/978-981-32-9244-4_28

gas distribution map by multi-chamber e-nose in a room gave a reference result [12]. It's interesting to compound some algorithms, which have their own advantages, into a new algorithm fitting for whole process of searching for underground gas source. This paper applies computer simulation to study how to measure and map the gas concentration of a single underground source landfill. In the simulation work, a virtual mobile robot with a gas sensor mounted on it is guided by a defined algorithm and measure concentration at the same time.

The rest of the paper is organized as follows. The advantages and disadvantages of the classic Hexagon-path Algorithm and several improved methods are discussed in Sect. 2. Then the functions and properties that a new algorithm should have are induced in Sect. 3. Also in the same section the paper suggests an implementable compound algorithm blending separated advantages and verifies the effectiveness and progressiveness of the new compound algorithm. In Sect. 4 the gas concentration field is mapped by means of the sequence of concentration values. Section 5 presents the conclusions.

2 Literature Review

2.1 Classic Algorithm

The Hexagon-path algorithm [2, 3] is the classic algorithm. It presents strong stability, some astringency (Figs. 1 and 2) and strong anti-interference (Fig. 3), though its convergence radius is a little too larger [5] (Fig. 2). When the length of steps is too long, the mileage might be too large. On the contrary, if the length is too short, traveling may even be stopped on the way (Fig. 4). *This classic algorithm is taken as the basic one* and all below methods are based on it.

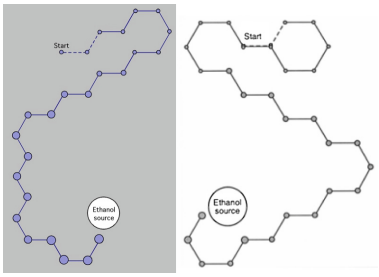


Fig. 1. Good stability and astringency of classic algorithm.

2.2 Method Based on Behavior Features and Concentration Difference

This method [5] varies length of step. Precisely speaking, it contains two sub-strategies: flexibly varying step size and surrounding walk with polycondensation. The robot changes its step size according to the difference of concentration values and traveling mode according to its behavior around the source. It partly solves the problem of

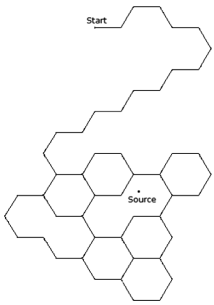


Fig. 2. Convergence radius is larger.

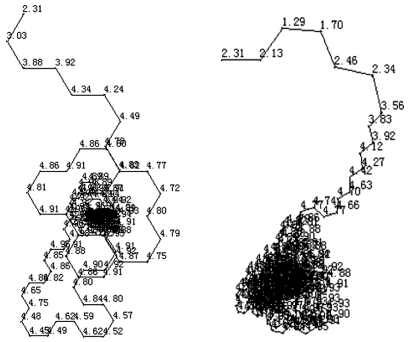
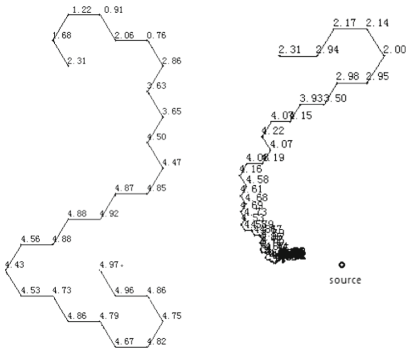


Fig. 3. Good astringency and anti-interference, small convergence radius.



a) too long step b) too short step

Fig. 4. Action of different step-size on search.

selection of length of steps. The method demonstrates small enough convergence radius which means *it has excellent performance in localizing the gas source*. And *it might be considered of traveling in middle stages*.

2.3 Method of Micro-steps Accumulation

This strategy admits that a gas sensor has sensitive threshold and the threshold acts on robot's behavior. It changes length of steps and angle of turns dynamically [6]. In other words, the method avoids the difficulty of artificially assigning step length and turn angle, and hands over the selection of step size to sensor's perception of environment. Thus *the selection of length of steps is partly resolved* and *the step sizes coincide with real distribution of concentration* (Fig. 5). The indicators of mileage and success rate are promoted significantly. The defection of the method is the application condition of formula of turn-angle calculation is too critic to judge. Once the formula is misused the robot will turn wrong direction or wrong angle (Fig. 6).

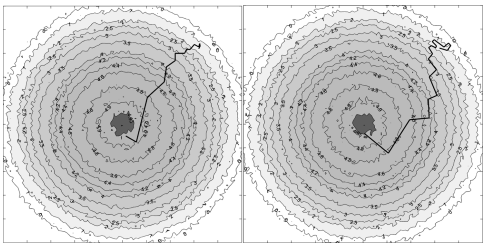


Fig. 5. Step sizes coincide with concentrations.

2.4 Method of Flexible Turn Angles

This method takes the principle of turn direction decision making which is suggested by classic algorithm, so it keeps the fundamental advantages of the classic one. It appends current concentration values and their first- and second-order differences to *promote the correctness of assessment of travel*, and *applies appropriate function to increase accuracy of turning angle*. The method usually has short mileage and high success rate (Fig. 7a), whereas weak anti-interference and robustness (Fig. 7b). The method depends on the parameters in algorithm, so the indicator of mileage fluctuates strongly when concentration with biggish noise.

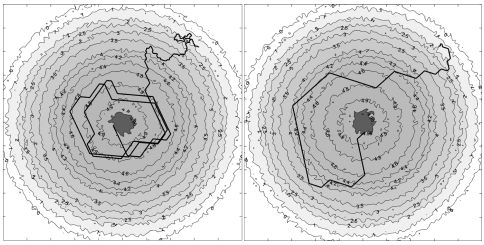
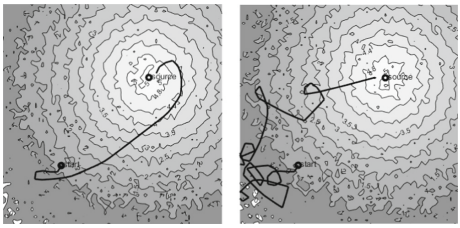


Fig. 6. Large convergence radius.



a) common paths b) bad anti-interference

Fig. 7. Results of flexible turning.

3 New Compound Algorithm

3.1 Algorithm's Structure and Its Model

Methods listed in previous section have their own advantages which appear in different stages of the search process. This hints that it is reasonable to treat the whole process of travel as several stages one after another and a compound algorithm which cascades the methods. In every stage the basic structure shown in Fig. 8 is adopted, where M presents search strategy and S is a double position switch. The switch is in initial position 1 and will turn to position 2 once the task is fulfilled.

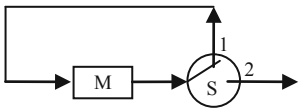


Fig. 8. Basic structure for stage search.

To realize the task of whole search process with a long journey, we propose a compound algorithm consisting of four stages and five switches whose complete model is shown in Fig. 9. Mode M_1 is used for the first stage, i.e., the initial stage of the process. Mode M_2 is used in half way stage. Modes M_3 and M_4 are fit for two different end stages, respectively.

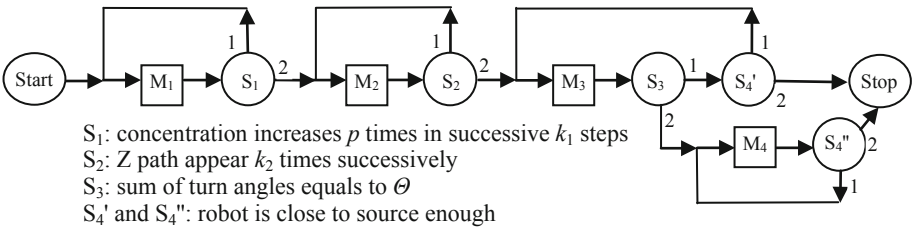


Fig. 9. Structure model of complete compound algorithm.

M_1 is a function which applied the method of micro-step accumulation [6] (Fig. 10), where N_{thresh} is sensor's sensitive threshold, $N(j)$ is the concentration value measured at stop j, Δl is the length of a micro-step.

```
repeat{
    wriggle straight Δl
 if |N(j+Δl)-N(j)|>N_thresh
 then
          jump out of loop
 else
          back to repeat
}
```

Fig. 10. Pseudocode of micro-step accumulation.

M_2 is a function which applied the method of dynamic variable step length [5] (Eq. 1), where l_n is the length of step n, k_1 is constant ($0 < k_1 < 1$).

$$l_n = \left(1 - k_l \cdot \frac{N(n) - N(n-1)}{N(n-1)}\right) l_{n-1} \tag{1}$$

M_3 is a function which applied the method of dynamic variable angle length [5, 13] (Eq. 2), where θ_n is the angle of step n, k_θ is constant ($0 < k_\theta < 1$).

$$\theta_n = \left(1 - k_\theta \cdot \frac{N(n) - N(n-1)}{N(n-1)}\right) \theta_{n-1} \tag{2}$$

M_4 is a function which applied the method of surrounding and polycondensation [5] (Eq. 3), where k_c is constant ($0 < k_c < 1$).

$$l_n = k_c . l_{n-1} \tag{3}$$

In brief, the algorithms discussed in last section are mixed in one compounded algorithm.

3.2 Simulation of Compound Algorithm

Computer simulation codes is programed in Matlab 7.13 and conducted in WINDOWS 10. Some of the critical observations made from an extensive study of the simulation results are summarized as follows:

(1) result after S_2 switching to position 2, i.e., stage M_2 being finished. The paths similar to those shown in Fig. 11 reach 78% of all paths. Other paths similar to those shown in Fig. 12, which are worse than that in Fig. 11, take about 8% of all paths. Simulation tells us that the best experience data is $k_1 = 7$, $p = 6$, $k_2 = 4$.

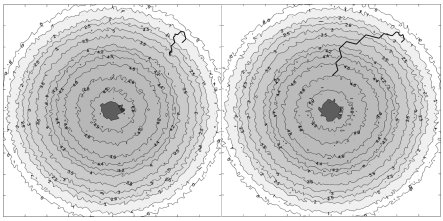


Fig. 11. Representative paths in on-way stage.

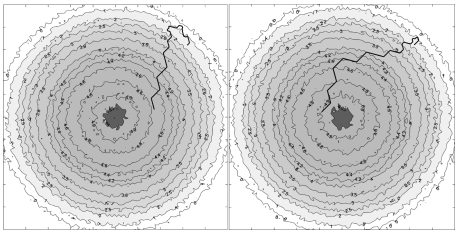


Fig. 12. Bad kinds of paths in on-way stage.

(2) result after stage M_3 being finished, i.e., S_3 keeping at position 1 and $S_4^{'}$ switching to position 2. Practically, paths shown in Fig. 11 must go the route $M_3 \rightarrow S_3 \rightarrow S_4^{'} \rightarrow$Stop (Fig. 13).

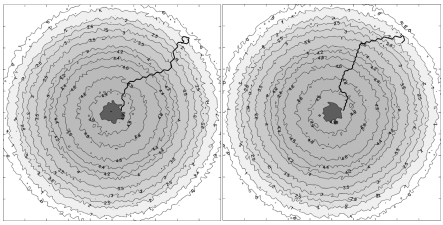


Fig. 13. Paths in end-stage "M3→S3→$S_4^{'}$→Stop".

(3) result after stage M_4 being finished, that is, S_3 switching to position 2 and $S_4^{''}$ switching to position 2. The paths shown in Fig. 12 must choose $M_3 \rightarrow S_3 \rightarrow M_3 \rightarrow S_4^{''} \rightarrow$Stop as their routes (Fig. 14).

This new compound algorithm is compared with the micro-step accumulation method in [6]. The simulation results of the compound algorithm and the accumulation method are listed in Table 1. It can be seen that the compound one brings out much less mileage, though it causes a little more steps and angles since the length of steps gets shorter in the stage of End. In one word, the cost of less increasing steps and angles is exchanged for large decreasing of mileage.

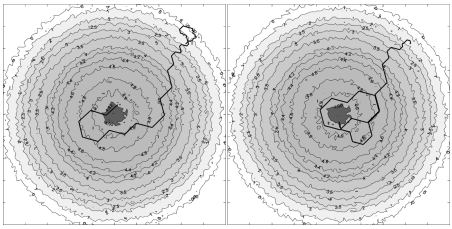


Fig. 14. Two paths in "M3→S3→M4→S_4''→Stop".

Table 1. Results of two algorithms. ($k_1 = 7$, $p = 6$, $k_2 = 4$, $k_l = 0.5$, $k_\theta = 0.9$, $\Theta = \pm 300°$, $k_c = 0.97$)

Algorithm	Average steps	Average mileage	Total angles	Success rate
Compound algorithm	32	26	1768	100%
Micro-step accumulation method	25	51	1689	100%

4 Mapping of Gas Field

Let step i starts at spot (x_i, y_i) and stops at spot (x_{i+1}, y_{i+1}). Thus a rectangle of concentration with above two spots as a pair of vertexes, can be noted as **Rec(*i*)**, that is $[x_i, y_i, x_{i+1}, y_{i+1}, A_i]$. The α_i presents the direction of the rectangle in 2-dimension plane. For **Rec(0)**, we let $A_0 = 0$. Thus for **Rec(*i*)**, we have $A_i = A_{i-1} + \alpha_i = \alpha_1 + \alpha_2 + \ldots + \alpha_{i-1} + \alpha_i$. When search is finished, the source is localized, the line segment connecting the Start and the End is definite. Then A_0 can be assigned as Θ that is the practical angle between the line segment and diagonal of **Rec(0)** (seeing Fig. 15). Consequently, the directions of all rectangles are assured.

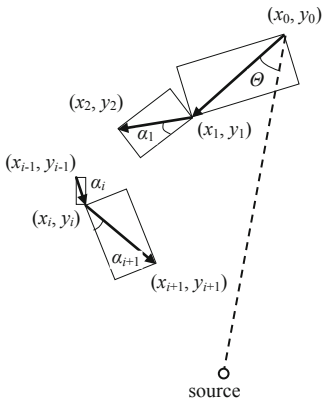


Fig. 15. Angle of concentration rectangles.

The source of gas is set after just three searches according to the three end spots. The center of the circumscribed circle of the three end spots is taken as the source.

Figure 16 demonstrates the sequence of concentration rectangles with angles being already rotated by Θ degrees. Figure 17a is the concentration field generated by letting all rectangles revolves around the source. Figure 17b shows the given concentration field and the measured field. The former is the solid lines which mark mid-values of concentrations and the latter is the dashed lines which mark maximum values of concentrations. We calculate the position of gas source according to three end positions. Figures 17 and 18 show that the error of mapped field is acceptable. Even though the deviation of gas concentration is larger long away from the source than near the source, the error of measurement is almost constant. That is because the N_{thresh} is small (=0.2), the length of step in the first stage is small.

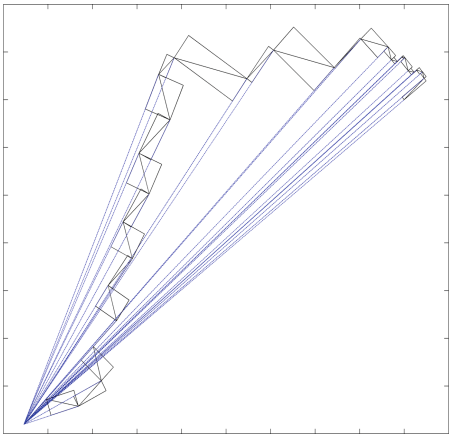
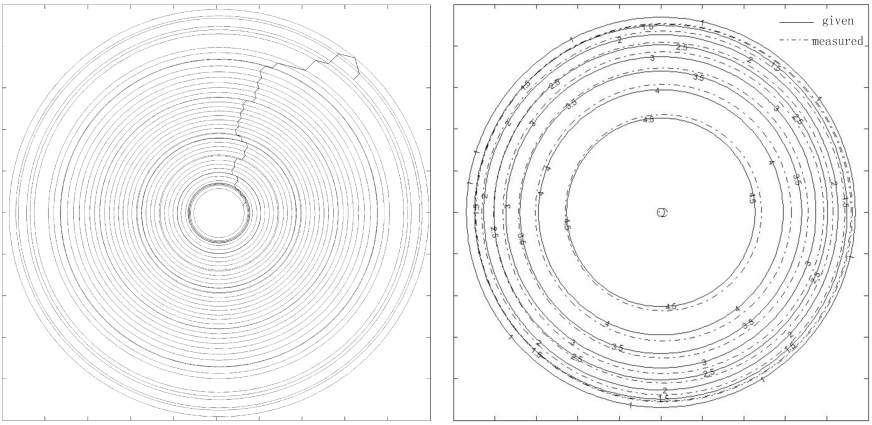


Fig. 16. Sequence of rectangles of concentration after rotation transformation.



a) a path and the measured field b) final given field and measured field

Fig. 17. Concentration field by measuring-mapping and given field.

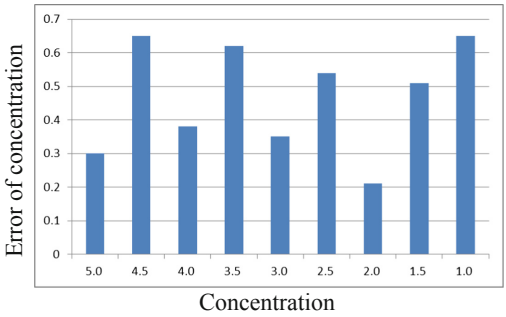


Fig. 18. Maximum errors between given and measured field.

5 Conclusions and Future Work

This paper has studied compounding several methods for achieving the whole process of gas source search with a relative long journey. The compound algorithm is verified feasible. The following results were validated by computer simulation: the most important indicator, mileage, is promoted remarkably, whereas the less important indicators such as average steps and total angles are a little worsen; the success rate keeps 100%; the method of using measured values to form small concentration rectangles, then rotating and montaging them to map whole gas concentration field is feasible; mapping is fast because it only needs three searches; the map has acceptable accuracy. The compound algorithm proposed can be applied in measuring and mapping gas concentration field formed by single underground source.

Our subsequent work will focus on the concentration field formed by multi underground gas sources. Gas sensors and sensors arrays, or sensors network may be applied in future work. Group of mobile robots should be tried to use simultaneously.

Acknowledgment. The authors would like to thank the anonymous reviewers for their invaluable comments, which improved this manuscript. This research is supported by National Natural Science Foundation of China (Grant No. 61672371).

References

1. Abaci, S., Edwards, J.S.: Migration of methane and leachate from landfill. In: International Conference on Environmental Management, Geo-Water and Engineering Aspects, pp. 347–350, Taylor & Francis, Inc., Wollongong (1992)
2. Russell, R.A.: Chemical source location and the RoboMole project. In: Australasian Conference on Robotics and Automation, CD-ROM Proceedings, pp. 1–6 (2003). ISBN 0-9587583-5-2
3. Russell, R.A.: Robotic location of underground chemical sources. Robotica **22**(1), 109–115 (2004)

4. Russell, R. A.: Locating Underground Chemical Sources by tracking Chemical Gradients in 3 Dimensions. In: International Conference on Intelligent Robots and Systems, IEEE/RSJ, pp. 325–330, 28 September–2 October 2004
5. Wang, J., Ji, J.L., Chen, W.D.: Behavior-feature-based dynamic variant step-size odor search algorithm for mobile robot. J. Syst. Simul. **24**(17), 5427–5430, 5435 (2009)
6. Wang, J., Chen, Z., Lei, Z.: Olfaction threshold based thought of micro-steps accumulation and its application on odor source search. IEEE Sens. J. **18**(17), 7294–7304 (2018)
7. Chen, Z., Wang, J.: Underground odor source detection based on a variation of lower organism search behavior. IEEE Sens. J. **17**(18), 5963–5970 (2017)
8. Lu, G.D., Zhang, M.L.: A biological-inspired gas/odor source localization strategy for mobile robot. J. Hebei Univ. Technol. **39**(5), 48–52 (2010)
9. Xie, Y.C., Pan, X.Y., Wang, J.: An odor source searching strategy on the base of detected concentration polygon. Comput. Sci. **24**(S1), 112–114 (2015)
10. Xie, H.T., Zhang, C.Z., Xu, F., et al.: Concentration levels and spatial distribution of heavy metals in soil surrounding a municipal solid waste incineration plant (Shenzhen). Environ. Sci. **35**(4), 1523–1530 (2014)
11. Li, X.X., Zhang, B., Wan, Z.M., et al.: Application of golden software Voxler in the investigation and health risk assessment for contaminated site. Sci. Technol. Eng. **17**(8), 317–323 (2017)
12. Gonzalez-Jimenez, J., Monroy, J.G., Blanco, J.L.: An improved olfaction sensor for mobile robotics. Sensors **11**, 6145–6164 (2011)
13. Wang, Y., Meng, Q.H., Li, T., et al.: Single-robot odor source localization in a ventilated indoor environment using simulated annealing algorithm. Robot **35**(3), 283–291 (2013)
14. Lay-Ekuakille, A., Trotta, A.: Predicting VOC concentration measurements: cognitive approach for sensor networks. IEEE Sens. J. **11**(11), 3023–3030 (2011)

Computing Service Scheme with Idle Virtual Machine Based on OpenStack

Jueun Jeon[1], Seungchul Kim[1], Gisung Yu[1], Hyun-Woo Kim[2], and Young-Sik Jeong[1(✉)]

[1] Department of Multimedia Engineering, Dongguk University, Seoul, Korea
{jry02l07, sc6742, gsyu, ysjeong}@dongguk.edu
[2] Department of Information Security, Baewha Women's University, Seoul, Korea
l0l63@baewha.ac.kr

Abstract. Virtual machine provides cloud computing services that offer computing resources to users through the Internet based on virtualization technology. Generally, supercomputing or grid computing has been used to process a large scale job in scientific, technology, and engineering application problems. Currently, services for large scale parallel processing through idle virtual machines in cloud computing are not provided. Previously, the utilization rate of computing resources in cloud computing was low when users do not use virtual machines anymore or for a long period of time since all the rights in relation to the use of virtual machine are given to users. This study proposes a scheme that increase resource utilization of idle virtual machines and process a large scale job through the idle virtual machines. Basically, idle virtual machines are identified based on virtual machines created through OpenStack, and idle virtual machine-computing service (IVM-CS) is proposed.

Keywords: Cloud computing service · OpenStack · Virtual machine · Computing resource · Distributed computing

1 Introduction

Generally, complex or data intensive problems are processed in parallel using super or grid computing. A study on large scale job processing using supercomputers or grid computing has focused on task partitions such as Cat Swarm Optimization, Particle Swarm Optimization, and K-means. However, an average operation ratio of network or server resources used in most institutions and companies is less than 20 to 30%, which is not fully utilized [1–4, 6, 7]. In addition, the importance on cloud computing that analyzes and processes a large scale of unstructured data has increased as much attention has been paid to the Fourth Industrial Revolution in Korea and overseas. The cloud computing services enable computing resources to be utilized efficiently as it integrates a number of physical resources into a single logical resource or divides a single physical resource into a number of logical resources to accommodate the user requirements based on virtualization technology [3, 5, 7–10]. Thus, studies on processing a large scale job in a distributed manner by utilizing cloud computing have

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 207–212, 2020.
https://doi.org/10.1007/978-981-32-9244-4_29

been conducted. And this paper proposes a virtual machine management at idle state to add computing service function by utilizing an idle virtual machine inside OpenStack, which is an open source cloud platform environment that supports multi platforms.

2 Related Works

Zhang [3] proposed a method to utilize idle virtual machines by using the garbage collector principle inside the OpenStack environment to provide computing resources. His method could prevent cloud's resource waste but did not provide computing services by utilizing idle virtual machines.

Praveenkumar [4] proposed a parallel data processing framework based on Nephele that managed computing resource dynamically. In this framework, virtual machines that were at an idle state for a long period of time were removed and transferred to other virtual machines that needed computing resources. However, it should solve the time problem required to remove virtual machines at an idle state for a long period of time and provide additional computing resources to other virtual machines.

Beloglazov [10] proposed OpenStack Neat, which was a framework to dynamically integrate virtual machines, by which low utilized virtual machines were migrated and over-loaded virtual machines were made off-load to other hosts. However, it should consider having virtual machines in a single host at the same time since conflict may occur if more than two global managers that decided the assignment of virtual machines were present simultaneously due to the service expansion.

3 Scheme and Design of IVM-CS

The idle virtual machine-computing service (IVM-CS) that provides cloud computing services by selecting idle virtual machines existed inside OpenStack is configured as shown in Fig. 1.

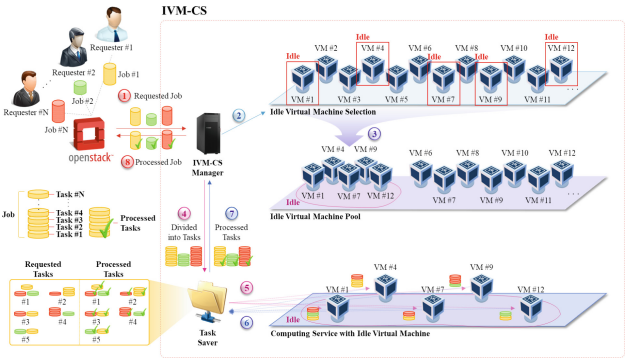


Fig. 1. Basic operation diagram of the IVM-CS

3.1 Idle Virtual Machine Selection

The idle virtual machine selection searches and selects virtual machines at an idle state at which users do not utilize the virtual machines. The virtual machine state can be classified into used and idle through the following four cases. The larger the case number is, the more the details of the idle state decision criteria are.

Case 1. When virtual machine is turned on, it is a used state and virtual machine is turned off, it is an idle state.
Case 2. When a virtual machine is turned on, and a user is logged and uses the virtual machine, it is a used state. However, when a user is logged out and does not use the virtual machine, it is determined to an idle state.
Case 3. When a virtual machine is turned on and a user is logged in the virtual machine, and the number of background processes is larger than zero, it is a used state. If the number of background processes is zero, it is an idle state.
Case 4. Considering Case 3, and additionally, a virtual machine use time log by day, week, month, and year is analyzed and the idle state of the virtual machine is predicted.

3.2 Idle Virtual Machine Pool

To provide computing services based on the idle virtual machines selected in the idle virtual machine selection in the above, static and dynamic information of the idle virtual machines is needed. The IVM-CS manager that manages virtual machines request static and dynamic information from idle virtual machines out of all virtual machines created in OpenStack.

The idle virtual machines and collected information are managed in the integrated idle virtual machine pool, and changing dynamic information can be reflected in the idle virtual machine pool immediately as the IVM-CS manager requests and collect the idle virtual machine information constantly.

3.3 Computing Service with Idle Virtual Machine

In this paper, the Sobel edge detection is provided as a computing service and the execution procedure is shown in Fig. 2.

Once the requester asks the task through the OpenStack dashboard, the IVM-CS manager divides the task uniformly based on the number of the selected idle virtual machines and task unit. The divided tasks by the IVM-CS manager are transferred to the task saver, in which each task is stored. The tasks processed by the idle virtual machines selected as a worker are transferred and stored in the task saver. The tasks stored in the task saver are transferred to the IVM-CS manager, and those tasks are merged into a single job. The IVM-CS manager displays the task processed results requested from the requester through the OpenStack dashboard.

This paper provides the IVM-CS function in addition to Horizon, Cinder, Nova, Swift, Neutron, Glance, and Keystone services that are basically provided by the OpenStack. The architecture of the OpenStack is shown in Fig. 3.

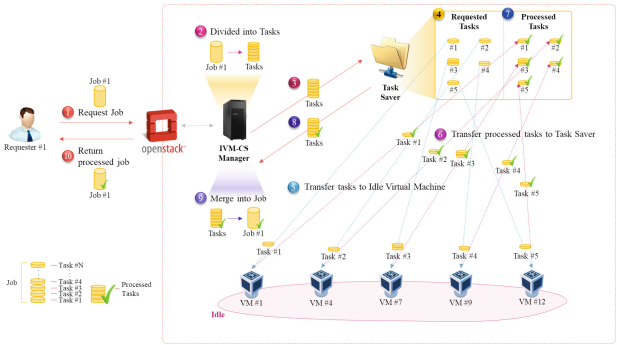


Fig. 2. Computing service control flow in IVM-CS

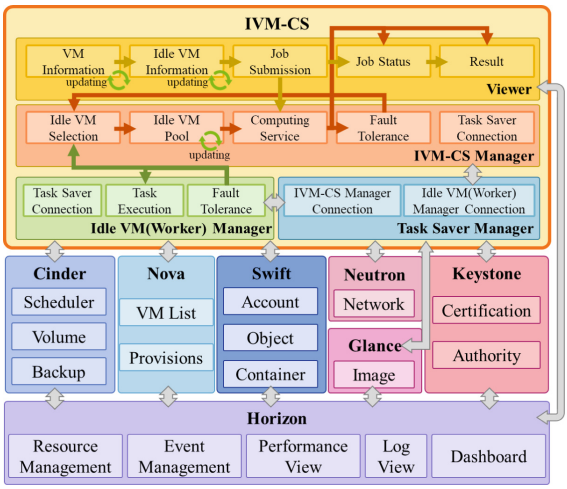


Fig. 3. IVM-CS architecture

IVM-CS largely consists of Viewer, IVM-CS Manager, Idle VM (Worker) Manager, and Task Saver Manager. Viewer is a visualized IVM-CS as a web format through horizon in OpenStack. It is composed of VM information that displays static and dynamic information of all created virtual machines, idle VM information that displays static and dynamic information of idle virtual machines, job submission that submits tasks, job status that displays the current processing task status, and result that displays the processed task results.

The IVM-CS manager consists of idle VM selection that selects idle virtual machines, idle VM pool that collects and manages static and dynamic information of the selected virtual machines, computing service that processes a large scale job in parallel by designating a worker out of the idle virtual machines, fault tolerance that resolves job losses or status change of the idle virtual machine into running, and task saver connection that is a network connection management function to store the

divided tasks in the task saver or merge the processed tasks. Computing service provides intermediate results in the job status of the viewer, and result provides the final results.

The idle VM (worker) manager consists of task saver connection that is a network connection function required to assign tasks from the task saver or store the processed tasks in the task saver, task execution that processes tasks assigned to idle virtual machines, and fault tolerance that resolves the status change and task loss in idle virtual machine occurred during task execution.

The task saver manager consists of IVM-CS manager connection and idle VM (worker) manager connection to share the tasks between IVM-CS manager and idle VM (worker) manager.

4 IVM-CS Implementation

The execution screen of the proposed IVM-CS is shown in Fig. 4. Figure 4-① checks the status information of the virtual machine. Figure 4-② shows the computing service to be executed, which is Sobel edge detection. Figure 4-③ selects the image that is processed as a task, and Fig. 4-④ designates the number of tasks distributed and processed in each of the workers. Figure 4-⑤ shows the running task's status, and Fig. 4-⑥ shows the result that merges completed process tasks into a single job.

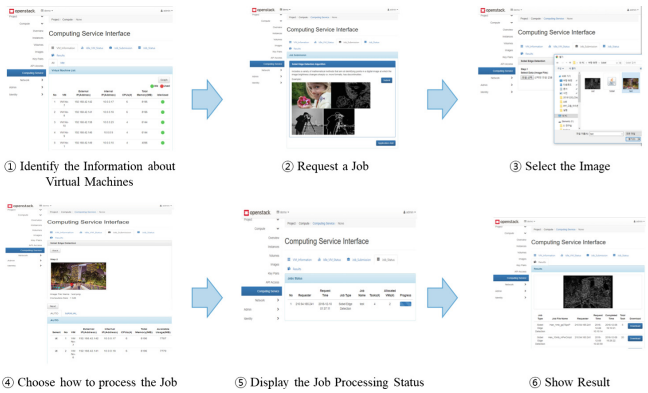


Fig. 4. Example of the execution procedure of Sobel edge detection by IVM-CS

5 Conclusions

This study proposes a computing service function by selecting and utilizing idle virtual machines that was not provided by existing OpenStack. The IVM-CS function can display the processed results requested from the requester as a web format through the OpenStack's dashboard, and define the status of the virtual machine. In addition, it divides and stores tasks through the IVM-CS manager and task saver and transfers the

tasks to idle virtual machine, and processes a large scale job in parallel by storing and merging the processed tasks.

As described in the above, it can minimize a waste of computing resource by increasing the utilization of computing resource through re-use of idle virtual machines and providing computing services via the IVM-CS function in OpenStack.

This study provided only the Sobel edge detection as a computing service by the proposed IVM-CS. However, a study on a method that provides not only Sobel edge detection, but also various computing services will be conducted for the future research.

Acknowledgments. This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No. 2018-0-00644, Linux Malware Dynamic Detection & Protection Solution on Embedded Device).

References

1. Kim, K., Jang, J., Hong, J.: Loan/redemption scheme for I/O performance improvement of virtual machine scheduler. Smart Media J. **5**(4), 18–25 (2016)
2. Rosenblum, M., Garfinkel, T.: Virtual machine monitors: current technology and future trends. IEEE Comput. **38**(5), 39–47 (2005)
3. Zhang, B., Al Dhuraibi, Y., Rouvoy, R., Paraiso, F., Seinturier, L.: CloudGC: recycling idle virtual machines in the cloud. In: IEEE International Conference on Cloud Engineering, pp. 105–115. IEEE Computer Society, Vancouver (2017)
4. Praveenkumar, V.P., Sujatha, D.N., Chinnasamy, R.: Efficient dynamic resource allocation using Nephele in a cloud environment. Int. J. Sci. Eng. Res. **3**(8), 1–5 (2012)
5. OpenStack Docs: Overview. https://docs.openstack.org/liberty/install-guide-ubuntu/overview.html. Accessed Jan 2019
6. Huh, J.-H., Seo, K.: Design and test bed experiments of server operation system using virtualization technology. Hum. Centric Comput. Inf. Sci. **8**(26), 1–21 (2016)
7. Kemchi, S., ZitouniEmail, A., Djoudi, M.: AMACE: agent based multi-criterions adaptation in cloud environment. Hum. Centric Comput. Inf. Sci. **6**(1), 1–28 (2018)
8. Ha, N., Kim, N.: Efficient flow table management scheme in SDN-based cloud computing networks. J. Inf. Process. Syst. **14**(1), 228–238 (2018)
9. Lim, J., HeonChang, Yu., Gil, J.-M.: An intelligent residual resource monitoring scheme in cloud computing environments. J. Inf. Process. Syst. **14**(6), 1480–1493 (2018)
10. Beloglazov, A., Buyya, R.: OpenStack neat: a framework for dynamic and energy-efficient consolidation of virtual machines in OpenStack clouds. Concurrency Comput. Pract. Experience **27**(5), 1310–1333 (2015)

N-Crop Based Image Division in Deep Learning with Medical Image

JuHyeon Lee, Dongho Lee, Yan Li, and ByeongSeok Shin(✉)

Department of Computer Engineering, INHA University, Incheon, Korea
leon6095@gmail.com, dongholabl3@gmail.com,
{leeyeon,bsshin}@inha.ac.kr

Abstract. In this paper, we propose an image division technique that can solve the problem of resolution reduction due to model structure and the lack of data caused by the characteristic of medical images. To verify this technique, we compared the performance of traditional full image learning and divided image learning. As a result, it is confirmed that the image division technique can proceed X-ray image deep learning more stable and is effective in predicting tuberculosis detection with higher accuracy.

Keywords: Deep learning · CNN · Medical image · Data augmentation

1 Introduction

Recently, there have been many studies using deep learning to diagnose diseases [1]. The performance of deep learning can be improved with complementing problems using techniques that match the characteristics of the data and model used. Deep learning with X-ray image also has some problems due to its characteristics.

Convolutional Neural Network (CNN) [2] is a structure mainly used for deep learning using an image. To train model, input images must be resized because models using CNN structure mostly have a fixed input size. If the original image is of higher resolution, there must be more reduction of image resolution. And it also means a reduction in the accuracy of the model [3]. X-ray images are also lack due to legal problems and high cost as a medical image, which means that the dataset available for model learning becomes small. This also leads to a decrease in the accuracy of the model [4]. To improve the resolution problem, existing studies have used methods to improve the resolution of the original data by adding additional computation processes before model learning or by cutting only the important part [5]. And to improve the lack of dataset, they increased dataset through distortion such as rotation, symmetry, and so on [6, 7].

In this paper, unlike traditional methods, we propose an image division technique that can solve these two problems efficiently. Through this technique, the image is divided into several patches before model learning, resulting in improved image resolution and the effect of data augmentation. In addition, we proved that this technique works efficiently by comparing the performance between learning the full image and learning the divided image.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 213–218, 2020.
https://doi.org/10.1007/978-981-32-9244-4_30

Section 2 describes our proposed techniques and experimental methods, and Sect. 3 evaluates how our image division technique works on the learning of the model. Section 4 shows a conclusion.

2 Image Division Technique in Deep Learning

In general, for the original X-ray image has a size of about 2000 × 2000, if resize it to 224 × 224 and results in a quality drop about 8 times. However, resize to five 224 × 224 patches, resulting in 488 × 488 pieces of information that can be learned more than twice as much as learning with entire images. In addition, five times the number of images to be learned can lead to more stable learning. Therefore, the two problems of deep learning using X-ray images can be improved by the proposed image division technique (Fig. 1).

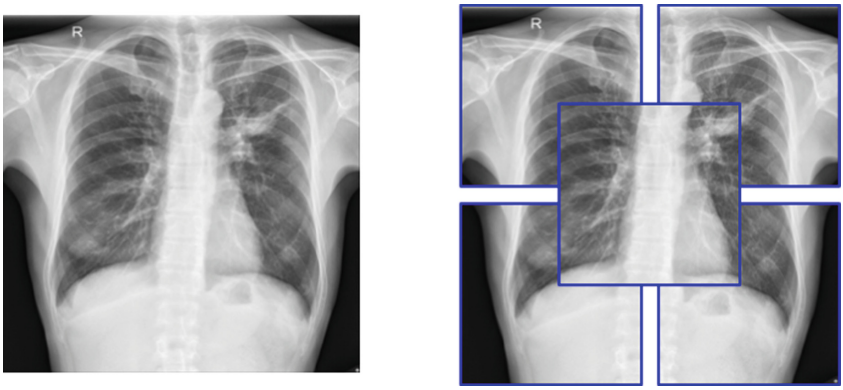


Fig. 1. The result of image division. Five patches are divided from the original image, one square form each from the four peaks of the image, and one partial square in the middle.

The experiment used X-ray image dataset provided by the U.S. National Library of Medicine. All images used were marked whether tuberculosis is existing. Dataset used in training is the Shenzhen Hospital X-ray Set (a total of 662) consist of X-ray images of 326 normal patients and 336 tuberculosis patients. And we used the Montgomery County X-ray Set consist of X-ray images of 80 normal patients and 58 tuberculosis patients to examine whether overfitting has occurred.

The preprocessing was conducted in two ways according to the experiment, the first is to learn the entire image, and the second is to divide the image into five patches. If the entire image was used, the image was resized to 224 × 224 size, and then the left and right reverses were performed with a 50% probability and normalized before learning. When using Image Division, the image was resized to 448 × 448 sizes, twice the size, and then divided into five patches size of 224 × 224 and normalized before learning.

The model used in the experiment is the flagship CNN model, DenseNet121 [8], and the initial weight is the weight which is initialized with ImageNet. The optimizer used is Adam optimizer and initial learning rate was set to 0.0001, adjusted by using ReduceLROnPlateau Scheduler to reduce the learning rate when the loss is stagnant. Loss function used is Binary Cross Entropy Loss because the problem is a binary classification problem whether tuberculosis is existing.

$$\text{loss}(O_m, T_m) = \sum_{k=1}^{n} \frac{-(O_m^k \log T_m^k + (1 - O_m^k) \log(1 - T_m^k))}{n} \tag{1}$$

Both train and validation were carried out with batch size 4 and proceed 100 epochs. In the case of general learning, train calculated the loss of the model extracted from the image and performed the optimization process, but in case of using Image Division, the average of the output was calculated in five patches. Formula 1 is our loss calculation formula, n means the number of patches and O_m means mth output, T_m means mth target. Validation also performed the output as the average of the loss, as with the training. After both learning and validation, evaluation examined the over-fitting of the model by including the Montgomery dataset which was not included in the training.

A Guided-Backpropagation technique [9], one of the visualization techniques [10], was used to determine which areas the model focuses on the image. Guided-Backpropagation is a technique that executes Backpropagation with Deconvolution [11], which performs a backward by setting the node which has negative gradient value or negative input value to 0.

3 Results

In both cases, the measurement range of loss initially tends to be wide. A closer look at the graph shows that if the entire image has been learned, it does not converge into a constant loss even after 100 epoch learning. However, when learning a divided image, the range of loss is smaller than in the case of the entire image, and one can see the convergence from about 20th epoch. It presents the Image Division technique can lead to stable learning, which means saving computing power by reducing learning time.

Visualizing the entire image learning shows that the model is focusing on the white tissue area in the upper right corner mainly. On the other hand, if the model was learned divided images, you can see that the model is looking at not only the upper-right part but also the other part. This shows that the Image Division helps the model make an accurate prediction by increasing the features that the model can see (Figs. 2, 3 and Table 1).

The AUROC may be increased from 0.75 to 0.85 for China dataset, and from 0.68 to 0.78 for the dataset include Montgomery set after Image Division, respectively. In addition, it is indicated that overfitting did not occur as the AUROC fell when the Montgomery dataset was added, and it also means that the training was progressed

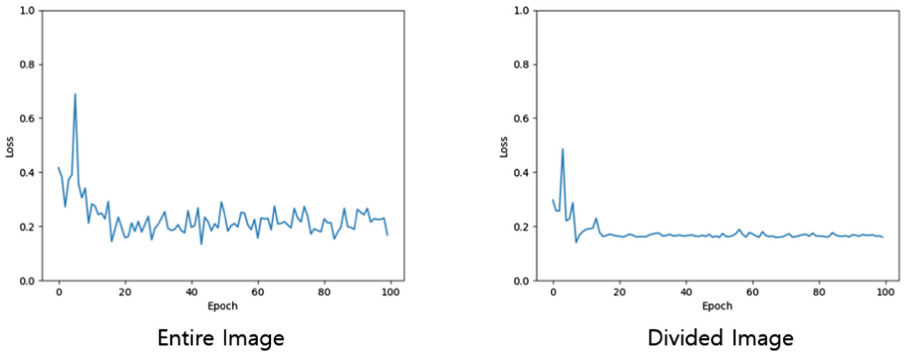


Fig. 2. Validation Loss Graph. Left graph is validation loss when the model is trained with entire images. And the right graph is validation loss when the model is trained with divided images.

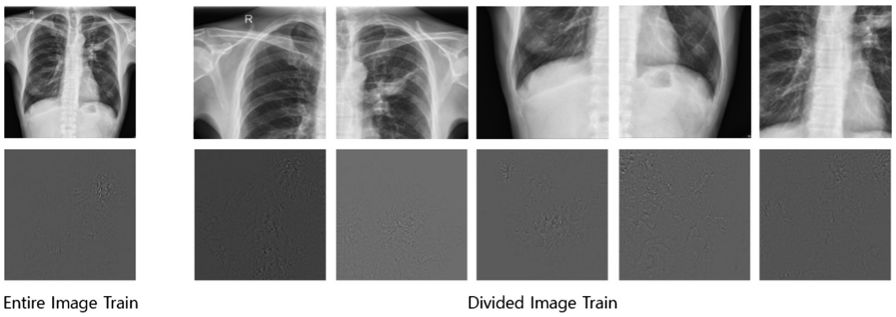


Fig. 3. Visualization results. Figures are visualized images of which parts the model focuses on when it learns the entire image and the divided image through Guided-Backpropagation technique.

Table 1. AUROC of each case. The case which was trained with divided images showing higher AUROC.

Dataset	AUROC (Entire image)	AUROC (Divided image)
China	0.75	0.85
China + Montgomery	0.68	0.78

normally. Finally, we can know the Image Division technique can model learned more features and increase prediction accuracy by progress train stably (Fig. 4).

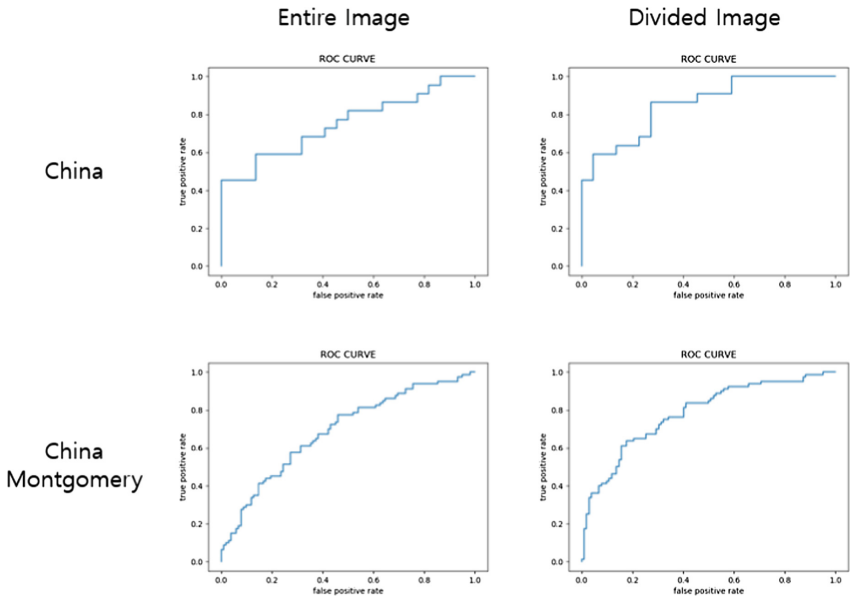


Fig. 4. ROC curves, when the entire image has been trained and when it has been trained with divided images. The upper two graphs are those examined by only China dataset, while the lower two graphs are those examined with the dataset include Montgomery set.

4 Conclusion

Through the results, we were able to obtain a more stable loss graph when trained divided images than when trained entire images. Visualization results also represent, as expected, capture more features than when we trained the entire images, and with these results, we were able to achieve a saving of computing power and improved prediction accuracy. This means that the Image Division technique is an efficient technique in deep learning using X-ray images. Hopefully, these techniques will help advance disease-diagnostic technology, allowing more people to find diseases quickly and eventually reduce patients suffering from diseases around the world.

References

1. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M.P., Ng, A.Y.: CheXNet: radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv:1711.05225 (2017)
2. Lecun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. In: Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324 (1998)
3. Dodge, S., Karam, L.: Understanding how image quality affects deep neural networks. In: Eighth International Conference on Quality of Multimedia Experience (QoMEX), Lisbon, pp. 1–6 (2016)

4. Soekhoe, D., van der Putten, P., Plaat, A.: On the impact of data set size in transfer learning using deep neural networks. In: Bostrom, H., Knobbe, A., Soares, C., Papapetrou, P. (eds.) Advances in Intelligent Data Analysis XV, IDA 2016. Lecture Notes in Computer Science, vol. 9897. Springer, Cham (2016)
5. Zheng, L., Zhao, Y., Wang, S., Wang, J., Tian, Q.: Good practice in CNN feature transfer. arXiv:1604.00133 (2016)
6. Perez, L., Wang, J.: The effectiveness of data augmentation in image classification using deep learning. arXiv:1712.04621 (2017)
7. Takahashi, N., Gygli, M., Pfister, B., Van Gool, L.: Deep convolutional neural networks and data augmentation for acoustic event detection. arXiv:1604.07160 (2016)
8. Huang, G., Liu, Z., van der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4700–4708 (2017)
9. Springenberg, J.T., Dosovitskiy, A., Brox, T., Riedmiller, M.: Striving for simplicity: the all convolutional net. arXiv:1412.6806 (2014)
10. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Grad-CAM: visual explanations from deep networks via gradient-based localization. In: IEEE International Conference on Computer Vision (ICCV), Venice, pp. 618–626 (2017)
11. Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional networks. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) Computer Vision? ECCV 2014. Lecture Notes in Computer Science, vol. 8689. Springer, Cham (2014)

An Empirical Study on Continuance Using Intention of OTT Apps with Young Generation

Thi-Thanh-Quy Tran[1], Quoc-Tuan Tran[2], and Hoanh-Su Le[3(✉)]

[1] Economic and Business Administration Faculty, Dalat University, Da Lat, Vietnam
tranquy2ll@gmail.com
[2] Candidate of School of Graduate Studies, Asia e-University, Kuala Lumpur, Malaysia
reb.arch.design@gmail.com
[3] E-commerce Department, University of Economics and Law Vietnam National University Ho Chi Minh City, Ho Chi Minh City, Vietnam
sulh@uel.edu.vn, lehoanhsu@gmail.com

Abstract. Today, Over-The-Top (OTT) apps have become more popular in many countries in the world. Although the development of OTT apps has contributed a lot of benefits to human life, its growth also has many disadvantages for traditional media suppliers. Therefore, with the goal of providing suppliers and managers with many methods to improve their apps services, this study examined factors that influence on continuance using intention OTT apps with the young generation in Vietnam. A survey of 288 Vietnamese consumers in the age from 18 to 35 was conducted to test the theoretical model and structural equation modeling was employed in the subsequent data analysis. The result indicate that four factors including satisfaction, habit, emotional trust, perceived usefulness directly influence on using intention with OTT apps. In addition, the mediation effects of satisfaction were found between perceived usefulness, perceived enjoyment, socialites and intention to use. And perceived usefulness and perceived enjoyment also have indirectly impacts on using intention. Based on these findings, several implications and recommendations are discussed.

1 Introduction

OTT (Over-the-top) is the name of media apps or services delivered over the Internet [16]. In recently, the emergence of OTT made traditional media business model disrupted [4]. In Vietnam, according to the Vietnam Mobile User Behavior [2], having 91% consumers used OTT apps every day to network and chat. In addition, in the Vietnam Moblie Report Q2 [3], Facebook Messenger (37%) and Zalo (30%) stood at the highest positions on top 5 popular message apps in Viet Nam next to Skype (21%), Viber (18%), Wechat (13%)… The growth of OTT apps led to a significant degree in traditional telecommunication services as well as had a bad impact on Short Message Services (SMS) [18]. Thus, to hold old consumers and attract new users, managers need to understand their consumer continuance using behavior. Therefore, this study

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 219–229, 2020.
https://doi.org/10.1007/978-981-32-9244-4_31

was conducted to achieve the following subjects: (1) determine factors influence on continue intention using OTT apps of the young generation; (2) evaluate the level impact of elements on intention to continuance usage OTT apps.

2 Previous Researches

Some previous researches related continuance using intention such as "Exploring the influential factors in continuance usage of mobile social Apps: Satisfaction, habit, and customer value perspectives" [13]. "The mediating effects of habit on continuance intention" [1]. "Continued usage intention of multifunctional friend networking services: A test of a dual-process model using Facebook" [10]. "The effects of cognitive and emotional trust in mobile payment adoption: a trust transfer perspective" [11].

2.1 Satisfaction

Satisfaction on continuance usage is defined as an overall awareness of consumers on using social apps and play an important role on adopting and keep a loyal relationship with consumers [13–20]. Besides, the satisfaction of the user was a reliable prediction on continuance using following a research on IS (Information System) and marketing [24–28]. In addition, satisfaction is also one of factors impact on continue usage [10]. Thus:

Hypothesis 1: Satisfaction positively affects continuance intention.

2.2 Habit

The habit of using social apps is understood as times that users tend to automatically access on social apps without thinking [15]. A number of previous studies have shown that when customer use these information technology devices regularly and automatically, some models used habit to explain the continuance using information technology [17]. Besides, habit is also one of factors impact on continuance usage [1–10]. Thus:

Hypothesis 2: Habit positively affects continuance intention.

2.3 Perceived Usefulness

Perceived usefulness is defined as the level of awareness that an individual believes using a particular service or system will improve his or her performance [8]. Besides, satisfaction was led from cognitive interpretation and related processes of using the product or service following the "two appraisals" satisfaction model [19]. In addition, perceived usefulness makes an increase customer trust and influence on their satisfaction [10]. Thus:

Hypothesis 3a: Perceived usefulness positively affects satisfaction.
Perceived usefulness is determined awareness of individual to action to get particular rewards. The previous experiences play an important role in building continuance

intention [9]. In addition, TAM (technology acceptance model) said that if individuals believe on the positive results of the information technology, they will accept it [8], that make them tend to use it regularly [13]. Thus:

Hypothesis 3b: Perceived usefulness positively affects on habit.
Moreover, perceived usefulness is as an expectation when consumers continue use IS that was shown on the post-acceptance model of IS continuance on explaining the intention of the IS user [13]. In addition, other research also indicates that use increases the trust of the customer and influences on the intention to continue to use [10]. Thus:

Hypothesis 3c: Perceived usefulness positively affects continuance intention.

2.4 Perceived Enjoyment

Perceived enjoyment is defined as pleasure from using technology [7]. Besides, an emotional response of an individual using social apps is the satisfaction of the user. Consequently, enjoyment affects user satisfaction by using social apps [13]. Thus:

Hypothesis 4a: Perceived enjoyment positively affects on satisfaction.
The feeling of enjoyment has an increase on the tendency to repeat of customer's actions, and this action occurs as a habit when they communicate by using social applications [13]. Thus:

Hypothesis 4b: Perceived enjoyment positively affects on habit.
In some studies of IS, Intrinsic motivation from perceived enjoyment has an important effect in shaping intentions using [26]. In addition, the value of enjoyment led to the causes of intention to use the product or service, following the consumer behavior literature [5]. Thus:

Hypothesis 4c: Perceived enjoyment positively affects continuance using intention.

2.5 Social Ties

Social influences have an important position on the intention to use products and services [30–35]. In addition, generally young generation usually are approved by others to use mobile or new devices thus social influence has a huge effect on them than older people [22]. Thus:

Hypothesis 5a: Social ties positively affects on satisfaction.
Daily using of a social application will become stable and may even reduce the use of other social applications if it becomes popular among group users [20]. When using familiar social applications, individuals will not spend much time thinking about using it and will probably use it regularly [13].

Hypothesis 5b: Social ties positively effect on habit.
Social factors have a strong influence on an individual's intentions that have been empirically proven [14]. Social relations play an important role in shaping the user's relationship market behavior and have a positive influence on the intention of a service [29]. Thus:

Hypothesis 5c: Social ties positively affect continuance intention.

2.6 Emotional Trust

Emotional trust in web payments refers to the feeling of comfort and safety of the customer about web payment services [21]. It is very important to build trust and to make the intention to continue to use for consumers. Emotional trust has a stronger impact than cognitive beliefs on the consumer and has a direct and indirect impact on the intention to use [11]. Besides, the consumer will accept the product if they have a high level of emotional confidence in their emotional beliefs. Thus:

Hypothesis 6: Emotional trust positively affect continuance intention.
From discussion above, following research model was proposed (Fig. 1).

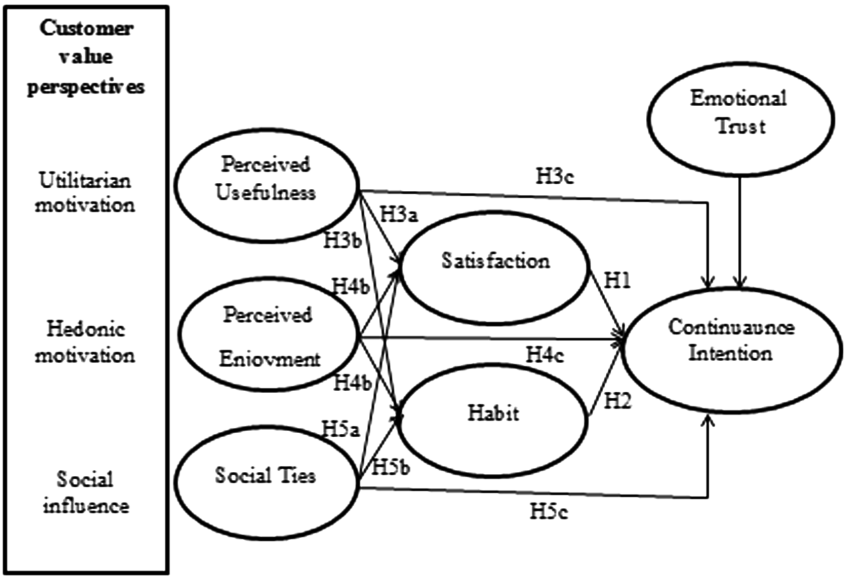


Fig. 1. Research model

3 Methods

The steps used to test the research model is a preliminary and formal study. Preliminary studies include qualitative research through interviews conducted on samples of size n = 20. The research was done by formal quantitative methods with the structured questionnaire based on a 5-point Likert scale, convenient sampling method with sample size n = 288 (126 paper, 162 interviews online). A survey was performed with Vietnamese consumers in the age from 18 to 35 living in Vietnam. Collected data used to test the reliability by Cronbach Alpha, the validity of scale by EFA, testing theoretical models with CFA and SEM.

Encode scales: perceived usefulness (PU), perceived enjoyment (EN), social ties (ST), continuance intention (CI), emotional trust (EMP), satisfaction (SA), habit (HA).

Descriptive Statistics

Samples included 133 (46.2%) female customers and 155 (53.8%) male customers in total. In terms of education, 231 (80.2%) customers had a university degree, 16 (5.6%) were under university degree, 41 (14.2%) had the postgraduate degree. In terms of income, 100 (34.7%) of customers earned less than 4 million VND/month, 96 (33.3%) earned from 4 to 9 million VND/month, 53 (19.1%) %) earned from 10 to 20 million VND/month and 37 (12.8%) earned more than 20 million VND/month. In terms of occupation, 64 (22.2%) customers were office workers, 84 (29.2%) were students, 39 (13.5%) were business, 95 (33%) were the government worker. Some other characteristics of respondents are summarized in Fig. 2 as follows. In the terms of used OTT apps, Facebook Messenger with 252 (87.5%) stood the first top of quantity user, follow by Zalo 145 (50.4%), Viber 67 (23.3%) and others (Skype, Whatsapp, Line, Wechat, Twitter) (Figs. 3 and 4).

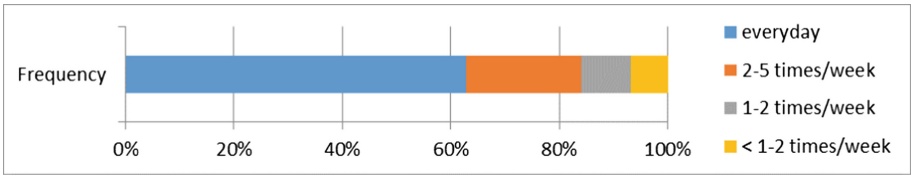


Fig. 2. Frequency of using OTT apps of respondents

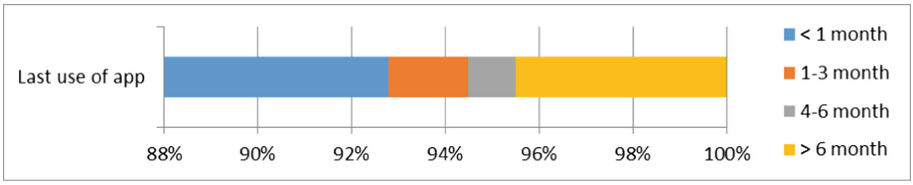


Fig. 3. Last time using OTT apps of respondents

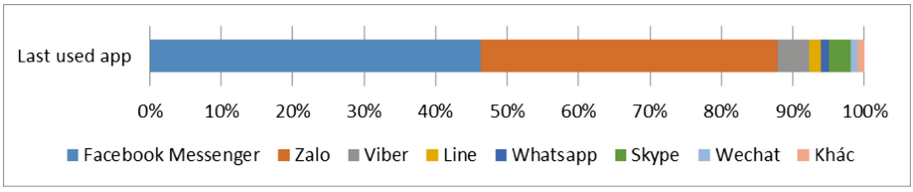


Fig. 4. Last used OTT apps of respondents

4 Results

4.1 Scale Validation

Scales were evaluated by using Cronbach's alpha reliability and EFA factor analysis with 288 respondents. Cronbach's alpha reliability test results have three variables EN08 (0.246), HA20 (0.233) and EMP27 (0.144) respectively for the perceived enjoyment, habit and emotional trust scale excluded because the Corrected Item-Total Correlation was less than 0.3. After removing the three variables from scales, scales reached the reliability with Cronbach's alpha higher 0.8. After that, two variables PU03 (0,328) and ST13 (0,395) having a weight of less than 0.5 were rejected from the results of the EFA (Principal Axis Factoring with Promax rotation). After removing these two variables, the result of the last analysis EFA show seven factors extracted in the Eigenvalue ($1.155 \leq 1$), KMO (KaiserMeyer-Olkin) ($0.852 > 0.5$), average variance extracted ($65.275\% > 50\%$) [23].

The CFA results show that the saturated model matches the market data: Chi-square = 373.237, df (degree of freedom) = 278, p = 0.000, and Chi-square statistics CMIN/df = 1.343 < 2. Indicators TLI (Tucker-Lewis Index) = 0.973, and CFI (Comparative Fit Index) = 0.977 were all greater than 0.9. And RSMEA (RootMean Square Error of Approximation) was 0.035 (less than 0.08). RSMEA (RootMean Square Error of Approximation) was 0.064 (less than 0.08). Therefore, this model as suitable for market data. The CFA weights were high (≥ 0.690) and statistically significant ($p < 0.05$; Table 1). Moreover, the above results also show that the scales meet the requirements for Composite Reliability ($CR \geq 0.859$) and Average Variance Extracted ($AVE \geq 0.577$). Thus, the scales of research concepts meet the reliability and validity requirements [12].

Table 1. Normalized CFA weights of the observed variables.

				p
Perceived usefulness: CR = 0.905; AVE = 0.705				
PU01	<—	Perceived usefulness	0.803	***
PU02	<—	Perceived usefulness	0.869	***
PU04	<—	Perceived usefulness	0.828	***
PU05	<—	Perceived usefulness	0.857	***
Perceived enjoyment: CR = 0.880; AVE = 0.648				
EN06	<—	Perceived enjoyment	0.731	***
EN07	<—	Perceived enjoyment	0.824	***
EN09	<—	Perceived enjoyment	0.821	***
EN10	<—	Perceived enjoyment	0.839	***
Social ties: CR = 0.869; AVE = 0.611				
ST11	<—	Social ties	0.731	***
ST12	<—	Social ties	0.704	***
ST14	<—	Social ties	0.847	***
ST15	<—	Social ties	0.835	***

(*continued*)

Table 1. (*continued*)

				p
Satisfaction: CR = 0.859; AVE = 0.577				
SA16	<—	Satisfaction	0.814	***
SA17	<—	Satisfaction	0.760	***
SA18	<—	Satisfaction	0.704	***
SA19	<—	Satisfaction	0.757	***
Habit: CR = 0.939; AVE = 0.775				
HA21	<—	Habit	0.847	***
HA22	<—	Habit	0.898	***
HA23	<—	Habit	0.895	***
Continuance intention: CR = 0.861; AVE = 0.582				
CI28	<—	Continuance intention	0.799	***
CI29	<—	Continuance intention	0.799	***
CI30	<—	Continuance intention	0.758	***
CI31	<—	Continuance intention	0.690	***
Emotional trust: CR = 0.905; AVE = 0.639				
EMP24	<—	Emotional trust	0.765	***
EMP25	<—	Emotional trust	0.802	***
EMP26	<—	Emotional trust	0.830	***

Besides, this result confirms the uni-dimensionality and convergent validity of the scales. Correlation coefficients between the concepts and their Standard Error (SE) show that these correlations are different from the unit, confirming discriminant validity between concepts [12].

4.2 SEM Results

Structural equation modeling (SEM) was used to test the theoretical model and hypotheses. The SEM results show that the theoretical model is consistent with market data: Chi–square = 507,531, df = 287, p = 0.000, and CMIN/df = 1.768 < 2. Indicators TLI = 0.940, CFI = 0.947 are all greater than 0.9 (reasonable fit), and RSMEA was 0.052 which is less than 0.08 [12].

SEM estimates of the theoretical model show that 9 in 12 relationships hypothesized in the model are statistically significant with a reliability of 95%, so that 9 hypotheses in the model are acceptable. Table 2 shows that the variables that affect directly on continuance usage OTT apps: satisfaction (β = 0.338; p = 0.003), habit (β = 0.271; p = 0.002), emotional trust (β = 0.210; p = 0.003), perceived usefulness (β = 0.320; p = 0.001) Besides, the variables that affect indirectly on continuance usage OTT apps through satisfaction: perceived usefulness (β = 0.163; p = 0.047), perceived enjoyment (β = 0.271; p = 0.001), social ties (β = 0.215; p = 0.013). And the variables that affect indirectly on continuance usage OTT apps through habit: perceived usefulness (β = 0.185; p = 0.013), perceived enjoyment (β = 0.209; p = 0.003).

Table 2. Hypothesis testing.

Relationship			(β)	P	Result
Usefulness	—>	Habit	0.185	0.013	Support H3b
Usefulness	—>	Satisfaction	0.163	0.047	Support H3a
Enjoyment	—>	Habit	0.209	0.003	Support H4b
Enjoyment	—>	Satisfaction	0.271	0.001	Support H4a
Social ties	—>	Satisfaction	0.215	0.013	Support H5a
Social ties	—>	Habit	0.072	0.310	Not support H5b
Habit	—>	Continuance intention	0.271	0.002	Support H2
Satisfaction	—>	Continuance intention	0.338	0.003	Support H1
Emotional trust	—>	Continuance intention	0.210	0.003	Support H6
Enjoyment	—>	Continuance intention	0.095	0.196	Not support H4c
Usefulness	—>	Continuance intention	0.320	0.001	Support H3c
Social ties	—>	Continuance intention	0.061	0.333	Not support H5c

Table 3 shows that perceived usefulness has the strongest impact (βtotal = 0.668) on continuance intention, followed by perceived enjoyment (βtotal = 0.48), satisfaction (βtotal = 0.338), habit (βtotal = 0.271), social ties (βtotal = 0.215) and finally, emotional (βtotal = 0.21). In summary, this table illustrates the direct and indirect impact of continuance intention using OTT apps of customers in Vietnam, Vietnam.

Table 3. Direct, indirect, and total impact of different factors on continuance intention

	Perceived usefulness	Perceived enjoyment	Social ties	Satisfaction	Habit	Emotional trust
Indirect effect	0.348	0.48	0.215	0	0	0
Direct effect	0.320			0.338	0.271	0.210
Total effect	0.668	0.48	0.215	0.338	0.271	0.21

5 Conclusions and Management Implications

The result shows that satisfaction is both one of the most significant influences and the mediate influence factor of use, enjoyment and social relations to continue intention using OTT apps. This means that consumers will be a pleasure with using this app when it meets their expectations by its useful and fun that make they want to recommend to other. In addition, if OTT apps help people feel relaxed, relieve stress and maintain good social relationships with friends, that will also enhance their satisfaction and intend to continue to use OTT apps in the future. In addition to satisfaction, the results show that habit is both a direct influence and a mediating factor of useful, perceived enjoyment to intend to continue using the OTT apps of young people. Consumers have the habit to use old, popular and familiar messaging apps, hence they regularly use OTT apps every day as well as these apps also help them to solve job so

fast, entertainment, simple, easy that improve their ability to manage their lives better. Besides, in terms of direct impacts on the intention to continue to use the OTT apps, emotional trust is one of the factors that have a direct impact. More and more people use the OTT apps because of not only its perceived usefulness but also their feelings comfort and security when using.

Moreover, the results show that perceived usefulness is the direct and indirect impact of satisfaction and habits on the continued use of the OTT apps. Previous research no direct effect on the intention to continue using the app (with $\beta = 0.01$ and $p > 0.05$) [13]. However, the results of this study show a direct effect of the intention to continue to use (with $\beta = 0.320$ and $p = 0.001$ less than 0.05). This is showed in Bhattacherjee's study in 2001 [6]. Nowadays there are many useful OTT apps that not only help users manage their personal lives better, save time in communication with friends and customers but also make them feel comfortable as well as create habits of use OTT apps and intend to continue using them in future. In three perspectives on customer value for perceived usefulness, enjoyment, and social relationships that affect satisfaction and habits, perceived enjoyment is the most influential directly and indirectly on intention to continue to use the OTT apps. In fact, besides chat function, OTT apps also help users feel happy, relax and relieve stressful when using. From there, users will be satisfied and have a habit of using the OTT apps as well as intend to use it in the future. Furthermore, the results show that social ties are one of the factors that affect user satisfaction when using the OTT apps. In fact, OTT apps are not just a bridge for connecting people, communicating with friends and family, but also helping people solve many problems in life such as jobs, consulting, maintaining relationships with customers without high cost. Besides, they are also simple and easy to use that enhance the number of users. This makes customers feel satisfied. Subsequently, they will intend to continue using this OTT apps in the future. In addition, the indirect influence of social ties on the habits and direct influence of perceived enjoyment, social ties on intention to continue to use [13].

However, this results did not show the direct effect of social relations on habits ($\beta = 0.072$ and $p = 0.310 > 0.05$) as well as direct impact of perceived enjoyment ($\beta = 0, 095$ and $p = 0.196 > 0.05$), social ties ($\beta = 0.061$ and $p = 0.333 > 0.05$) to the intention to continue to use OTT app. This can begin differences in culture, preferences, economics, politics, etc., leading to differences in the continued use behavior of Vietnamese consumers compared to other countries in the world. Consequently, vendors must constantly explore the needs of consumers, improve the product more modern, simple, easy to use, convenient, function…to attract consumers. In summary, this model contributes to the theoretical system of behavioral use that specifically intends to continue to use of consumers. In this way, OTT apps providers can use this research model to build appropriate business strategies, marketing strategies and attract customers. As other studies, although the effort has been completed, the study is inevitably limited by the following: using a convenient sampling method, the data collected may have low reliability in terms of the sample, size in the study was limited to only 288 acres compared to tens of thousands of people aged 18–35 living in Vietnam that make the representativeness was not high. And the results will be better if

we can further study the direct and indirect impact of other factors on the user's continuance using intention to the OTT apps. These are the limitations of this research but also the direction for further research.

References

1. Amoroso, D., Lim, R.: The mediating effects of habit on continuance intention. Int. J. Inf. Manag. **37**, 693–702 (2017)
2. Appota: Vietnam Mobile User Behavior 2016 (2016). https://appota.com/report.html
3. Appota: Vietnam Moblie Report Q2 2017 (2017). https://appota.com/report.html
4. Barclay, C.: Is regulation the answer to the rise of over the top (OTT) services? An exploratory study of the caribbean market. IEEE Xplore (2015). https://ieeexplore.ieee.org/document/7383647/
5. Bauer, H.H., Barnes, S.J., Reichardt, T., Neumann, M.M.: Driving consumer acceptance of mobile marketing: a theoretical framework and empirical study. J. Electron. Commer. Res. **6**(3), 181–192 (2005)
6. Bhattacherjee, A.: Understanding information systems continuance: an expectation confirmation model. MIS Q. **25**(3), 351–370 (2001)
7. Brown, S.A., Venkatesh, V.: Model of adoption of technology in the household: a baseline model test and extension incorporating household life cycle. MIS Q. **29**(4), 399–426 (2005)
8. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q. **13**(3), 319–340 (1989)
9. Dorsch, M.J., Grove, S.J., Darden, W.R.: Consumer intentions to use a service category. J. Serv. Mark. **14**(2), 92–117 (2000)
10. Gwebum, K.L., Wang, J., Guo, L.: Continued usage intention of multifunctional friend networking services: a test of a dual-process model using Facebook. Decis. Support Syst. **67**, 66–77 (2014)
11. Gong, X., Zhang, K.K., Zhao, S., Lee, M.O.: The effects of cognitive and emotional trust on mobile payment adoption: a trust transfer perspective. In: Pacific Asia Conference on Information Systems (PACIS). AIS Electronic Library (AISeL) (2016). http://aisel.aisnet.org/pacis2016/
12. Hair Jr., J.F., Black, W.C., Babin, B.J., Anderson, R.E.: Multivariate Data Analysis, 7th edn. Pearson, USA (2014)
13. Hsiao, C., Chang, J., Tang, K.: Exploring the influential factors in continuance usage of mobile social apps: satisfaction, habit, and customer value perspectives. Telemat. Inform. **33**, 342–355 (2015)
14. Karaiskos, D.C., Drossos, D.A., Tsiaousis, A.S., Giaglis, G.M., Fouskas, K.G.: Affective and social determinants of mobile data services adoption. Behav. Inf. Technol. **31**(3), 209–219 (2012)
15. Kim, S.S., Malhotra, N.K., Narasimhan, S.: Research note—two competing perspectives on automatic use: a theoretical and empirical comparison. Inf. Syst. Res. **16**(4), 418–432 (2005)
16. Kokaram, A., Crinon, R., Catania, N.: OTT (Over the top). IEEE Xplore (2015). https://ieeexplore.ieee.org/document/7306504/
17. Limayem, M., Hirt, S.G., Cheung, M.K.: How habit limits the predictive power of intention: the case of information systems continuance. MIS Q. **31**(4), 705–737 (2007)
18. Lutilsky, I.D., Ivíc, M.: Influence of OTT service providers on croatian telecommunication market. IEEE Xplore (2016). https://ieeexplore.ieee.org/abstract/document/7522233/

19. Oliver, R.L., Desarbo, W.S.: Response determinants in satisfaction judgment. J. Consum. Res. **4**(14), 495–508 (1988)
20. Quan-Haase, A., Young, A.L.: Uses and gratifications of social media: a comparison of Facebook and instant messaging. Bull. Sci. Technol. Soc. **30**(5), 350–361 (2010)
21. Sun, H.: Sellers' trust and continued use of online marketplaces. J. Assoc. Inf. Syst. **11**(4), 182–211 (2010)
22. Smetana, J.G., Campione-Barr, N., Metzger, A.: Adolescent development in interpersonal and societal contexts. Annu. Rev. Psychol. **57**, 255–284 (2006)
23. Tho, N.D.: Methods of Scientific Research in Business. Social Labor Publishing House, Ho Chi Minh (2011)
24. Thong, J.Y.L., Hong, S.J., Tam, K.Y.: The effects of post-adoption beliefs on the expectation–confirmation model for information technology continuance. Int. J. Hum.-Comput. Stud. **9**(64), 799–810 (2006)
25. Turel, O., Serenko, A., Bontis, N.: User acceptance of wireless short messaging services: deconstructing perceived value. Inf. Manag. **44**(1), 63–73 (2010)
26. Van der Heijden, H.: User acceptance of hedonic information systems. MIS Q. **28**(4), 695–704 (2004)
27. Venkatesh, V., Thong, J.Y.L., Xu, X.: Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q. **36**(1), 157–178 (2012)
28. Wang, Y.J., Hernandez, M.D., Minor, M.S.: Web aesthetics effects on perceived online service quality and satisfaction in an e-tail environment: the moderating role of purchase task. J. Bus. Res. **63**(9), 935–942 (2010)
29. Woisetschläger, D.M., Lentz, P., Evanschitzky, H.: How habits, social ties, and economic switching barriers affect customer loyalty in contractual service settings. J. Bus. Res. **64**(8), 800–808 (2011)
30. Huh, J.-H.: PLC-based design of monitoring system for ICT-integrated vertical fish farm. Hum.-Centric Comput. Inf. Sci. **7**(20), 1–19 (2017)
31. Huh, J.-H.: Smart Grid Test Bed Using OPNET and Power Line Communication. Advances in Computer and Electrical Engineering, pp. 1–425. IGI Global, USA (2017)
32. Huh, J.-H., Seo, K.: RUDP design and implementation using OPNET simulation. In: Computer Science and its Applications, LNEE, pp. 913–919. Springer (2015)
33. Ngu, H., Huh, J.-H.: B+-tree construction on massive data with Hadoop. Clust. Comput. 1–11 (2017)
34. Huh, J.-H.: Server operation and virtualization to save energy and cost in future sustainable computing. Sustainability **10**, 1–20 (2018)
35. Huh, J.-H., Kim, T.-J.: A location-based mobile health care facility search system for senior citizens. J. Supercomput. **75**, 1–18 (2018)

The Trend of New and Renewable Energies in the Socialist Republic of Vietnam and the Possibility of Introducing Korean Smart Grid

Thi-Hong Nguyen(✉)

Coordinator of VK (Vietnam-Korea)
Technological Exchange Seminar at Thanh Hoa,
Thanh Hoa, Socialist Republic of Vietnam
thuhong2903@hotmail.com

Abstract. It is considered that the Socialist Republic of Vietnam (Vietnam) has an ideal environment for introducing and spreading Smart Grids. Although power demands in Vietnam are increasing rapidly and continuously, the infrastructure for smart grids is still underdeveloped so that the demands have not been satisfied yet, and the same can be said for the other South-Eastern countries. Most of their thermoelectric power plants are generating power by using diesel generators and the hydroelectric power generation is another means of securing power. It seems that the governments in this region prefer a small-scale power system linked to new and renewable energies rather than supporting some large-scale power generation projects due to their regional characteristics being consisted of a number of islands where power infrastructures are inadequate. The demands for an independent power system are being demanded in the countries with many islands such as Indonesia, Malaysia or Thailand such that it is essential to find new export-oriented businesses in these areas and push ahead with the construction of test beds for smart grids. In this regard, this study considers the possibility of introducing a Korean model smart grid in Vietnam while considering their situation and trend of new and renewable energies.

Keywords: Renewable energies · Smart Grid · Smart grid service · Micro Grid

1 Introduction

While the power demands are continuously increasing in the Socialist Republic of Vietnam (Vietnam), they have not been satisfied due to the lack of adequate infrastructure for the construction of the Smart Grids.

Smart Grid is a Power Grid which pursues a high-level reliability and stability through effective and intelligent power management using its own communication network. The efficient use of power can be achieved with a cutting-edge Smart Grid technology which optimizes the energy efficiency by grafting the ICT technology on to existing Power Grid [1, 2].

LG Display at Hanoi, (former), Hanoi, Socialist Republic of Vietnam.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 230–238, 2020.
https://doi.org/10.1007/978-981-32-9244-4_32

The Vietnamese power market is managed and supervised by the Ministry of Industry and Trade (MoIT) who are in charge of establishing regulations, policies, development, strategies/plans pertaining to power and energies. The Vietnam Electricity Corp. (EVN) is a state-operated enterprise under the MoIT, monopolizing entire power distribution process as a sole power consumer or a distributor. In terms of power utilities, over 70% of the total power output in Vietnam is being produced by EVN (61.2%), Petrovietnam Oil & Gas group (PVN, 11.5%), Vietnam National Coal & Mineral Industries Group (Vinacomin, 4.6%), etc. [2016 Annual Report by EVN]. Table 1 is showing the current power supply flow in Vietnam.

Table 1. The current power supply flow

Power source	Production	Purchase & distribution	Consumption
Hydroelectric (42.1%)	*EVN (61%)	*Monopolized by EVN	*Residential (35.1%)
*Thermal (Coal: 34%, Gas: 17.4%)	*PVN (11.5%)	- EVN NPT	*Manufacturing/construction (53.7%)
	*Vinacomin (4.6%)	- 5 affiliated power corps	*Agricultural (1.6%)
New & renewable energies (5.4%)	*Others	(Hanoi, Ho Chi Minh, Nothern, Central, Southern)	*Commercial (5.3%)
			*Others (4.3%)

Even though the statistics vary depending on the organizations, the MoIT has estimated that the annual power production capacity is approximately 42,762 MW (as of 2016, including import) and the power demands are consistently increasing due to the rapid industrialization and urbanization. According to the MoIT, the rate of increase in annual power sales between 2010 and 2015 was over 10% and also, under the Power Development Plan IIV (2016), it was forecasted that the same rate would reach up to 10.7%.

Although the power demands will continuously rise every year, there are some concerns that demands would not be met in the future due to insufficient supply so that the Vietnamese government is planning to focus on the development of thermal power generation and gradually increase the power production based on new & renewable energies.

Smart Grid is a next-generation Power Grid which facilitates efficient use of energy through real-time information exchange between the electricity supplier and users based on the ICT-integrated Power Grid [3–6].

The Smart Grid is an excellent alternative to such a system having an advantage of minimizing such an energy waste by flexibly controlling the power supply by allowing the power company to check the real-time power consumption level while the users are able to use the power by automatically checking the time zones to which cheap rates are applied through the exchange of information between them. In short, the power is produced and supplied according to the power usage, minimizing energy waste [7–11].

All the things in a Smart Grid environment communicate with each other, even with humans through the intermediaries such as Smart Phones, computers, or other communication devices – expanding the domain of communications [12–14].

2 The Characteristics of Vietnamese Power Industry and Introducing Korean Smart Grid

One of the major characteristics of the power industry in Vietnam is that the current major sources of their power generations are hydroelectric and coal-fired power plants, occupying over 70% of entire power output. Specifically, the power generation sources are divided into water (42.1%), coal (34%), gas (17.4%), new & renewable energies (5.4%), and import (1.4%), being highly biased in favor of two aforementioned power generations. Meanwhile, according to the 7th Power Development Plan revised in March 2016, the power facility capacity was to be consistently increased up to 64.9% by 2025. One of the major issues of the hydroelectric power generation in Vietnam has been varied power production levels between seasons due to weather conditions. For example, from 2015 to 2016, there was a setback in power generations due to low water levels at the dams caused by El Nino. (Figure 1) is showing the proportion of the power output at each power generation source. The hydroelectric generation is strong in the northern areas compared to coal-fired generation in the central areas and gas-fired generation in the southern areas.

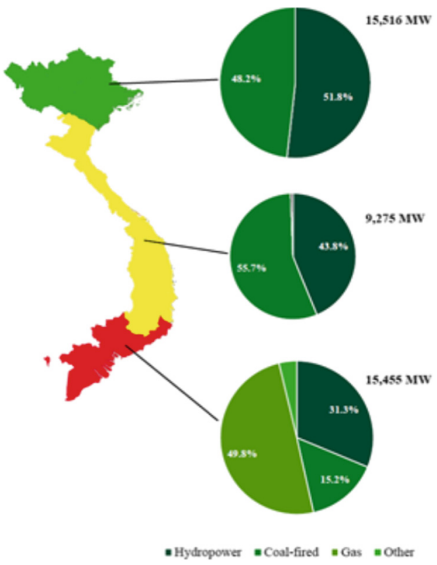


Fig. 1. The proportion of the power output at each power generation source in Vietnam.

The recent individual power generation capacities in these areas were 15,516 MW (Northern), 9,275 MW (Central), and 15,455 MW (Souther), being concentrated of hydroelectric, coal-fired, and gas-fired power generations, respectively (by EVN, from Jan. to Oct. 2016). The reason for the southern areas being biased in favor of gas-fired power generation (approx.. 50%) is because of their geographical factor – being close to the South China Sea where there are large deposits of oil and natural gas.

On March 18th, 2016, the 7th Power Development Plan was revised (Decision No. 428/QD-TTg) and announced. This new plan distinguished itself by emphasizing the development of new & renewable energy sources such as wind or solar power energy. Vietnam has then set up individual goals of increasing the power capacities based on these sources: 9.9%, 12.5%, and 21% increases by 2020, 2025, and 2030, respectively. The government's strong determination for such a strategy can be found in the fact that the contents of future projects and facility development plans have been concretely stipulated in the revision. However, the concerns expressed by some of the experts for the government's power/energy policy still remains as the plan relying heavily on fossil fuels could lead to increased environmental contamination. Thus, considering all the elements, the smart grid may shed light on such an issue.

Back to the revision made in March 2016, the plan aims to increase the proportion of coal-fired power production gradually from 34.4% (2015) to 49.3%, 55%, and 53.2% by 2020, 2025, and 2030, respectively while increasing the current power production capacity (approx.. 13,000 MW) to 55,300 MW by the end of 2030.

Some of the domestic and foreign experts expressed their concerns for such an excessive dependence on coal-fired power production and the vice-chairperson of the European Chamber of Commerce, Mr. Thomas Andreatta, warned by saying, "Excessive use of coal-based fuels will increase greenhouse gas emissions and lead to environmental contamination" and then suggested, "Vietnam should gradually reduce construction of coal-fired power plants." At the same time, the UNDP expected that if the plan proceeds accordingly, the number of early deaths due to environmental contamination would rise from 4,263 (2011) to 25,402 by 2030. (Figure 2) shows the proportion of power outputs by areas.

In addition, there are still many Southeast Asian countries with underdeveloped power systems, supplying power mainly with a hydroelectric generation or the diesel generators. Due to their geographical characteristics of being consisted of many islands, the governments prefer small-scale power systems that can be operated with new and renewable energies rather than supporting large-scale power generation projects. The demands for the independent power generation systems are high in Indonesia, Malaysia, Thailand, etc. so that the Korean power companies should consider finding a new export-oriented business opportunity there and proceed with construction of appropriate test beds. Meanwhile, Jeju-do and Ulleung-do are the islands of the Republic of Korea (ROK) and the former has experience in constructing and running a smart grid test bed while the latter possesses both smart grid and microgrid technologies with which they have built an energy-independent island.

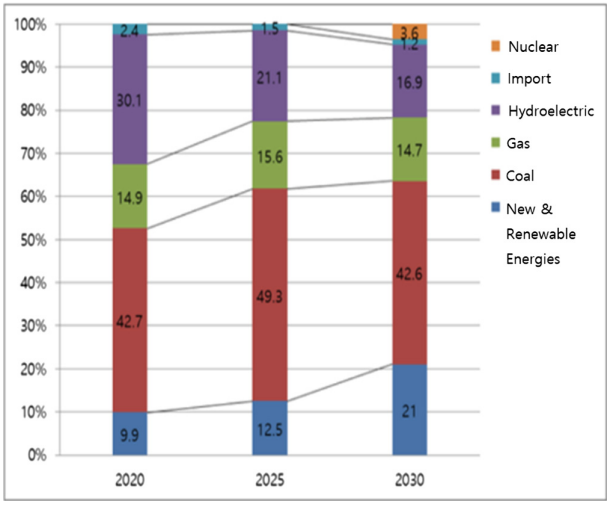


Fig. 2. The proportion of power outputs by areas.

(Figure 3) shows the trend on the Smart Grid systems in Republic of Korea. The Korean government has announced the 'Smart Grid Road Map' and plans to invest 2.75 billion won by 2030 hoping to vitalize relevant industries, starting with construction of a demonstrational complex in Jeju-Do. Their initial Smart Grid model had aimed to advance the element technologies but recently, they are seeking to invigorate the Smart Grid system business internationally by developing several promising

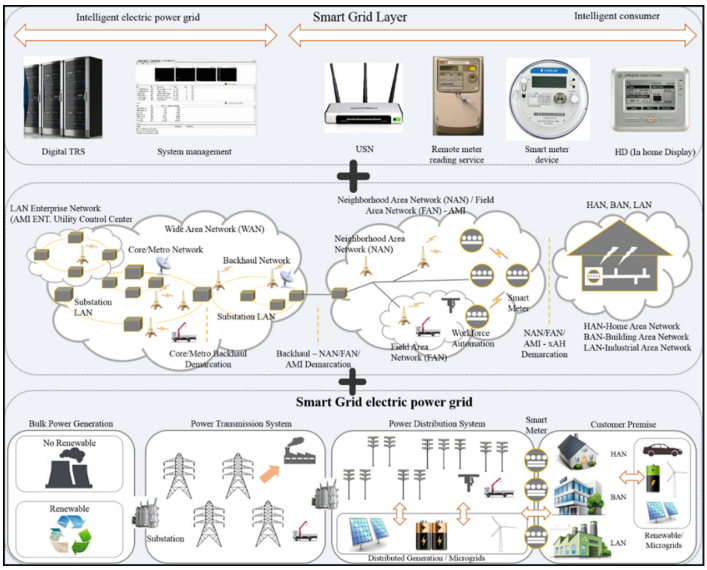


Fig. 3. A trend on the Smart Grid in Republic of Korea

business models and more sophisticated convergent systems. The purpose of such Smart Grid system is to intellectualize the Power Grid first and let the supplier of electric power to adjust the level of power supply and its rates and, on the part of consumers, to control their usage by understanding the price of electricity and their own usage patterns through the smart meters and internet. One of the typical methods used in the Smart Metering systems is a Cable TV (CATV) network-based remote metering method [15–19].

Owing to the rising interest in the Smart Grid and IoT technologies in ROK, major Korean companies such as Samsung, LG and KT are aggressively pushing forward with their smart home construction plans and many allied products have been recently introduced to the market [20–22]. The government of the ROK is making an effort to achieve the goal of Creation of Low Carbon Green Growth through Construction of Smart Grids by drawing and planning the "National Smart Grid Road Map" and "Smart Grid 2030," It is expected that 270 billion and 24.8 trillion Korean won will be invested by the government and private enterprises by the end of 2030, respectively (Ministry of Knowledge Economy, 2010). More importantly, they planned to develop the Smart Grid-related technologies by providing a test bed at the Smart Grid demonstration complex in Jeju. Considering such national-level interests and promising prospects for the technology, it will be fair to say that the analysis based on the mathematical model of micro grid, which is vital to the Smart Grid network, represents much significance [23–25]. Proposed platform [26–37] enables wireless communications based on WiFi and Zigbee, or wire communications that use the LAN system. Although the former has an advantage of not requiring any additional cable installations, there could be a problem of communication speed decline or lower reliability due to signal interferences at the final implementation process where all the electric appliances in the smart home are being interconnected and starting to communicate. The Power Line Communication (PLC) has been developed to complement this problem. This technology does not require additional wiring but instead, complements demerits of LAN- and wireless-based communications by forwarding data using existing 220 V household power line.

Meanwhile, Masayoshi Son elaborated his vision of the "Asia Super Grid" concept which was proposed in September 2011, and talked about the developments to date. Liu Zhenya, Chairman of Global Energy Interconnection Development and Cooperation Organization (GEIDCO) and Hwan-Eik Cho from Korea Electric Power Corporation (KEPCO) shared their ideas of international power grids that would connect Asia to the World [38]. Extending the technology of Smart Grid and Micro Grid will become a base technology for Super Grid.

3 Conclusion

A possibility of introducing a Korean-model smart grid into Vietnam in this study in addition to observing the situation and trend of use of new and renewable energies. The annual economic growth of Vietnam for the past 5 years was 6% in average but during the same period, the power demands had increased at a higher rate of 10% and it is still increasing rapidly to this day. Some domestic and foreign experts are concerned that

the demands would exceed the supply and inevitably lead to a power shortage. They are also worried that the thermal power generation-oriented power market could intensify environmental contamination. Following the Vietnamese government's power development plan, it is possible that the coal-fired power generation will occupy over 50% of entire power output and when added with the gas-fired power generation, the rate will reach up to 70%. As the current total power is between 13,000 and 15,000 MW, they plan to increase it to 55,300 MW by 2030, expecting that the number of fossil fuel power plants to be increased drastically from about 20 to maximum of 70. The many experts warn that air pollution will be intensified and the environmental costs can be much larger. Korean companies should pay more attention to the Vietnamese power market where a number of foreign-invested power companies are actively participating in. The Vietnamese government is implementing the BOT (Built Operate Transfer) system to cover the cost of infrastructure construction and has much interest in increasing the efficiency of the power transmission facility.

References

1. Farhangi, H.: The path of the smart grid. IEEE Power Energy Mag. **8**(1), 18–28 (2010)
2. Huh, J.-H., Seo, K.: Blockchain-based mobile fingerprint verification and automatic log-in platform for future computing. J. Supercomput. **75**, 1–17 (2018)
3. Huh, J.-H., Otgonchimeg, S., Seo, K.: Advanced metering infrastructure design and test bed experiment using intelligent agents: focusing on the PLC network base technology for Smart Grid system. J. Supercomput. **72**(5), 1862–1877 (2016)
4. Weaver, W.W., Krein, D.P.: Game-theoretic control of small-scale power systems. IEEE Trans. Power Deliv. **24**(3), 1560–1567 (2009)
5. Kasbekar, G.S., Sarkar, S.: Pricing games among interconnected microgrids. In: Proceedings of Power and Energy Society General Meeting, pp. 1–8. IEEE (2012)
6. Mohsenian-Rad, H., Wong, V.W.S., Jatskevich, J., Schober, R., Leon-Garcia, A.: Autonomous demand side management on game-theoretic energy consumption scheduling for the future smart grids. IEEE Trans. Smart Grid **1**(3), 320–331 (2010)
7. Bogdanov, D., Breyer, C.: North-East Asian super grid for 100% renewable energy supply: optimal mix of energy technologies for electricity, gas and heat supply options. Energy Convers. Manag. **112**, 176–190 (2016)
8. Feltes, J.W., Gemmell, B.D., Retzmann, D.: From smart grid to super grid: solutions with HVDC and FACTS for grid access of renewable energy sources. In: 2011 IEEE Power and Energy Society General Meeting. IEEE (2011)
9. Van Hertem, D., Ghandhari, M.: Multi-terminal VSC HVDC for the European supergrid: obstacles. Renew. Sustain. Energy Rev. **14**(9), 3156–3163 (2010)
10. Harris, A.: Smart grid thinking-[power super grid]. Eng. Technol. **4**(9), 46–49 (2009)
11. Aguayo, D., et al.: Link-level measurements from an 802.11 b mesh network. In: ACM SIGCOMM Computer Communication Review, vol. 34, no. 4. ACM (2004)
12. Huh, J.H., Seo, K.: Design and implementation of the basic technology for solitary senior citizen" s lonely death monitoring system using PLC. KMMS **18**(6), 742–752 (2015)
13. Lee, S., Huh, J.-H.: An effective security measures for nuclear power plant using big data analysis approach. J. Supercomput. 1–28 (2018)

14. Bychkovsky, V., et al.: A measurement study of vehicular internet access using in situ Wi-Fi networks. In: Proceedings of the 12th Annual International Conference on Mobile Computing and Networking. ACM (2006)
15. Petersson, N.A., Sjögreen, B.: Super-grid modeling of the elastic wave equation in semi-bounded domains. Commun. Comput. Phys. **16**(4), 913–955 (2014)
16. Andrew, B., Joachim, L., Anna, N.: Asia Pacific super grid–solar electricity generation, storage and distribution. Green **2**(4), 189–202 (2012)
17. Xu, Z., Dong, H., Huang, H.: Debates on ultra-high-voltage synchronous power grid: the future super grid in China? IET Gener. Transm. Distrib. **9**(8), 740–747 (2015)
18. Gordon, S.: Supergrid to the rescue [electricity supply security]. Power Eng. **20**(5), 30–33 (2006)
19. Dorri, A., Kanhere, S.S., Jurdak, R.: Towards an optimized blockchain for IoT. In: Proceedings of the Second International Conference on Internet-of-Things Design and Implementation, pp. 173–178. ACM (2017)
20. Kim, S.-K., Kwon, H.-T., Kim, Y.-K., Park, Y.-P., Keum, D.-W., Kim, U.-M.: A study on application method for automation solution using blockchain dapp platform. In: Park, J.H., Shen, H., Sung, Y., Tian, H. (eds.) PDCAT 2018. CCIS, vol. 931, pp. 444–458. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-5907-1_47
21. Huh, J.-H.: PLC-based design of monitoring system for ICT-integrated vertical fish farm. Hum.-Centric Comput. Inf. Sci. **7**(20), 1–19 (2017)
22. Huh, J.-H., Koh, T., Seo, K.: Design of a shipboard outside communication network and the test bed using PLC: for the workers' safety management during ship-building process. In: Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication, ACM IMCOM 2016, pp. 1–6 (2016)
23. Tran, B.: Mesh network personal emergency response appliance. U.S. Patent No. 7,733,224. USA, 8 June 2010
24. Talla, V., et al.: Powering the next billion devices with Wi-Fi. In: Proceedings of the 11th ACM Conference on Emerging Networking Experiments and Technologies. ACM (2015)
25. Bitar, E., Khargonekar, P.P., Poolla, K.: Systems and control opportunities in the integration of renewable energy into the smart grid. IFAC Proc. **44**(1), 4927–4932 (2011)
26. Huh, J.-H.: Implementation of lightweight intrusion detection model for security of smart green house and vertical farm. Int. J. Distrib. Sens. Netw. **14**(4), 1–11 (2018)
27. Kim, J., Park, H.I.: Policy directions for the smart grid in Korea. IEEE Power Energy Mag. **9**(1), 40–49 (2011)
28. Huh, J.H., Lee, D.G., Seo, K.: Design and implementation of the basic technology for realtime smart metering system using power line communication for smart grid. In: Advances in Computer Science and Ubiquitous Computing, pp. 663-669. Springer, Singapore (2015)
29. Song, I.K., Yun, S.Y., Kwon, S.C., Kwak, N.H.: Design of smart distribution management system for obtaining real-time security analysis and predictive operation in Korea. IEEE Trans. Smart Grid **4**(1), 375–382 (2013)
30. Huh, J.-H., Seo, K.: Hybrid advanced metering infrastructure design for micro grid using the game theory model. Int. J. Softw. Eng. Appl. 9(9), 257–268 (2015)
31. Kim, H., Baek, S., Park, E., Chang, H.J.: Optimal green energy management in Jeju, South Korea–on-grid and off-grid electrification. Renew. Energy **69**, 123–133 (2014)
32. Huh, J.-H.: PLC-integrated sensing technology in mountain regions for drone landing sites: focusing on software technology. Sensors **18**(8), 2693 (2018)

33. Ngar-yin, M., van der Vleuten, J.M., Chi-man, I., Ronald, H.: Governing the transition of socio-technical systems: a case study of the development of smart grids in Korea. Energy Policy **45**, 133–141 (2012)
34. Huh, J.-H.: Smart Grid Test Bed Using OPNET and Power Line Communication. IGI Global, Hershey (2017)
35. Mah, D., van der Vleuten, J.M., Ip, J.C.M., Hills, P.: Governing the transition of socio-technical systems: a case study of the development of smart grids in Korea. In: Smart Grid Applications and Developments, pp. 259-277. Springer, London (2014)
36. Huh, J.-H., Seo, K.: PLC-based smart grid home network system design and implementation using OPNET simulation. J. Multimedia Inf. Syst. **1**(2), 111–118 (2014)
37. Kim, S.K., Huh, J.H.: A study on the improvement of smart grid security performance and blockchain smart grid perspective. Energies **11**(8), 1973 (2018)
38. About "Asia Super Grid (ASG)". https://www.renewable-ei.org/en/asg/about/. Accessed 22 Feb 2019

Deep Learning Based IoT Re-authentication for Botnet Detection and Prevention

Mikail Mohammed Salim and Jong Hyuk Park(✉)

Department of Computer Science and Engineering, Seoul National University of Science and Technology (SeoulTech), Seoul 01811, Korea
{mikail, jhparkl}@seoultech.ac.kr

Abstract. IoT devices face a grave security threat from botnet attacks. These devices are known for their poor default authentication system due to being set on weak factory set passwords. Critical systems such as Healthcare and transportation can be jeopardized if hijacked. Using a bot, an attacker can use it to relinquish control from Smart city network administrators and users. In this paper, we present a Software-defined Deep learning based IoT Defense (SDID) mechanism which monitors and compares device historical traffic flow with current patterns to determine if a device is under an attack. Furthermore, to prevent false detection under flash-crowd events, the mechanism compares data with adjacent nodes to determine if the traffic flow is anomalous or not.

Keywords: Botnet · Cybersecurity · Internet of Things · Smart city · Deep Learning · Software defined networking

1 Introduction

IoT devices suffer from lack of basic authentication security protocols which motivates attackers to seize control over them to form a botnet. These devices have low computing power and battery resources, and if infected and disabled, it will result in critical systems failures such as in autonomous vehicles and health care.

There are multiple botnet defense solutions available that focus on detecting signature patterns of the attack, DNS information or anomalous traffic flow. These methods are unreliable as they are unable to identify new and evolving botnet attacks or detect suspicious traffic infused with normal traffic. Our detection and prevention model implements a machine learning based approach which can detect all unexpected patterns of a botnet attack on IoT devices.

2 Related Work

Addressing poor IoT device security, Bhunia et al.'s research [1] proposed a machine learning based SDN IoT defense framework. IoT device traffic is routed via switches and the SDN controller using the support vector machine algorithm detects suspicious traffic. Meidan et al.'s research [2] proposed a network-based botnet detection system (N-BaIoT) using deep autoencoders. It captures snapshots of historical data flow and

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 239–242, 2020.
https://doi.org/10.1007/978-981-32-9244-4_33

uses autoencoders to compress them. In the event, it is unable to reconstruct the snapshot, N-BaIoT indicates an anomalous behavior originating from the device. Mc Dermott et al.'s research [3] proposed using Bi-directional Long Short-Term Memory based Recurrent Neural Network based Deep Learning method (BLSTM-RNN) to detect botnet activity on consumer IoT devices. Among other features, it uses text recognition by converting it into integers using word embedding method to predict attack vectors.

Common drawbacks with the current research work are that they have no provision of detecting flash-crowd events which are common in IoT devices. Traffic flow pattern does always not remain constant. Secondly, they either suggest shutting down the device once detected as a bot, suggest rate limiting methods or quarantining the traffic which will have an adverse effect on critical and sensitive IoT devices which must always remain active.

3 Proposed Framework

In this section, we present the SDID based framework for IoT security with botnet detection and prevention. The proposed framework as shown in Fig. 1 is intended for a smart city-based environment consisting of different domains such as healthcare, vehicles and smart home. These domains have multiple IoT devices in place and provide services in collaboration with one another. The framework has the following advantages - (a) Addresses flash-crowd events with greater accuracy, and (b) should a switch be disabled, traffic to the controller will be a routed via nearby switch preventing failure of defence mechanism.

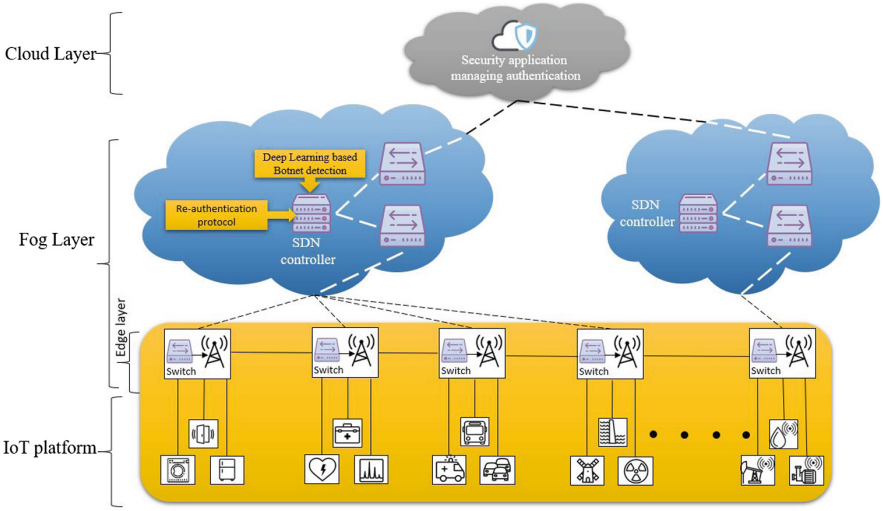


Fig. 1. SDID architecture

3.1 Framework Components

The framework consists of the following components: (a) Security application, (b) SDN controller, (c) SDN controlled switch, and (d) IoT devices.

The security application is present at the highest level of the proposed framework and is responsible for managing authentication on IoT devices. The application resets the authentication of the infected IoT device and informs the administrator. The SDN controller has master control over the network traffic and is responsible for setting traffic flow rules. The SDN controller upon detection of an infected device informs the security application of security breach to successfully prevent the device from being misused by the attacker. SDN managed switches are placed at the edge layer and consist of policies and rules to be followed. It is assumed that each smart city domain application is allocated an individual SDN compatible switch. IoT devices are used in different smart city domain application which ranges from home management, weather stations, smart grids and healthcare. These devices are directly connected with SDN switches. If a switch is under attack, traffic will be routed via nearby switches to the controller.

3.2 Operations of SDID Framework

IoT devices send regular traffic status to switches which monitor traffic flow patterns. The SDN controller component learns from traffic flow pattern features using Deep Learning algorithm as explained in Fig. 2 and detects any anomalous traffic based on known attack models. It is responsible for labelling the traffic either as normal or suspicious.

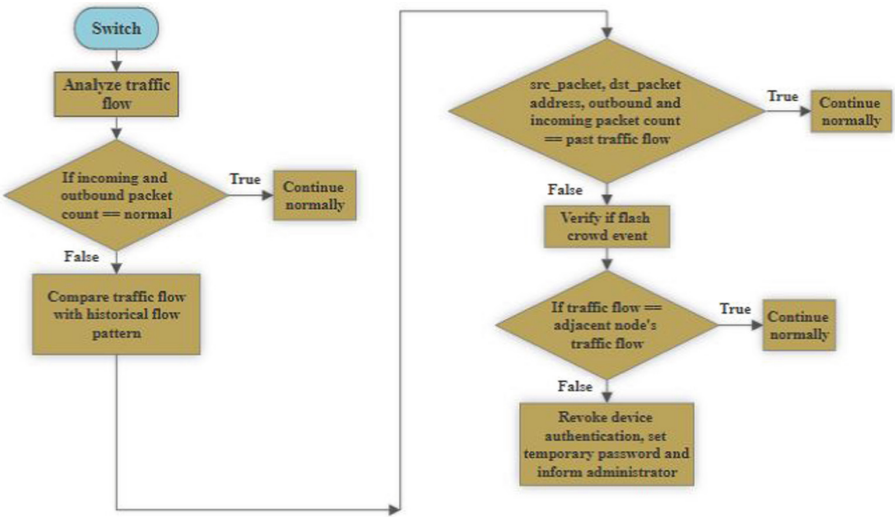


Fig. 2. Deep Learning process flow

To address the possibility of a flash-crowd event, traffic flow pattern from adjacent IoT devices will be checked which exist on the same network and are part of the same application. A smart home application would include devices such as thermostat, lighting and security. Smart Healthcare applications would include different devices connected to the patients responsible for providing medical care. If adjacent and supporting IoT devices show dissimilar data flow pattern, then the controller will mark the device as infected and inform the security application.

Prevention schemes such as Rate Limiting and partial or complete blocking of traffic flow results in hampering the functionality of devices. Normal resumption of these devices is critical for smart city applications where human life is at risk such as in healthcare. The security in the proposed method resets the authentication password and informs the administrator of the attack and password change. This step ensures services continue to resume normally while the attacker has lost control of the device. The administrator can reset the authentication protocol by assigning a new password.

4 Conclusion

IoT devices are prone to botnet attacks as they lack strong authentication passwords. Attackers using a brute force method manage to seize control over them. We addressed this weakness by proposing a botnet detection and prevention framework which detects anomalous traffic early before the device is misused or disabled by the attacker. We implement a machine learning technique, deep learning to detect abnormal traffic.

We are currently developing an implementation of the proposed framework and will publish results based on the practical deployment.

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

References

1. Bhunia, S.S., Gurusamy, M.: Dynamic attack detection and mitigation in IoT using SDN. In: 27th International Telecommunication Networks and Applications Conference, pp. 1–6. IEEE, Melbourne (2017)
2. Meidan, Y., Bohadana, M., Mathov, Y., Mirsky, Y., Shabtai, A., Breitenbacher, D., Elovici, Y.: N-BaIoT—network-based detection of IoT botnet attacks using deep autoencoders. IEEE Pervasive Comput. **17**, 12–22 (2018)
3. McDermott, C.D., Majdani, F., Petrovski, A.V.: Botnet detection in the Internet of Things using deep learning approaches. In: 2018 International Joint Conference on Neural Networks, pp. 1–8. IEEE, Rio de Janerio (2018)

FT2019

Fast Searching of Log Area in FAST FTL Using Log Area Indexing

Sang-Hyeok Yu and Tae-Sun Chung(✉)

Department of Software, Ajou University,
206, World cup-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do, Republic of Korea
{golazcc83, tschung}@ajou.ac.kr

Abstract. As flash memory is widely used for embedded devices, research on flash translation layer (FTL), a system software that reflects the pros and cons of flash memory, is steadily studied. FAST FTL, which stores the latest data in the form of a log, searches the log blocks to determine whether the data is updated. However, due to the low percentage of the latest data, it is inefficient to determine whether data is updated by exploring log blocks for all data. Therefore, this paper suggests a way to index information about log blocks to determine if data is up-to-date. The proposed method can improve read performance by indexing information about log blocks into a Binary Search Tree (BST) to quickly search the location of data that has been updated and not having to search log blocks.

Keywords: FAST · FTL · Flash memory · Indexing

1 Introduction

Flash memory is used for mobile embedded devices such as smartphones and MP3 players because it has a competitive price that can replace hard disks, is small and light compared to hard disks, and is strong in low power. However, because flash memory is structurally different from hard disk, it has characteristics such as 'Erase-before-write' that do not overwrite existing data to write new data, or limited number of records due to limitations in materials [1], and if used in an environment where existing hard disk is used, flash memory can degrade performance that flash memory can produce. Therefore, Flash Translation Layer (FTL) [2], which uses historical information to solve the problem of flash memory characteristics using additional system software for existing file systems, was studied.

FAST FTL is the most remarkable technique among the recently studied FTL methods [3]. It stores the latest version of data in the log area arbitrarily and searches the log area for each read and write operation. Since the latest version of the data is stored in the log area, when reading specific data, the log area is first searched to determine whether the data is up-to-date. However, as the size of the storage device becomes larger and the area of the log area becomes wider, the area to be searched becomes wider, and even if the data is not updated yet, the log area is searched firstly,

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 245–250, 2020.
https://doi.org/10.1007/978-981-32-9244-4_34

resulting in inefficient reading performance. In other studies, it has proposed a technique that can stochastically sort data that has not been updated using the counting bloom filter developed in other studies [4]. It is not different that the log area is searched in the same way even if the data is updated.

In this paper, we propose an indexing method of Physical Block Number (PBN) where the data is stored, the Logical Page Number (LPN), which is the original location of the updated data, and offset to Binary Search Tree (BST), which can detect both detecting whether data is updated and the data location.

2 Background

2.1 NAND Flash Memory

NAND flash memory consists of a page, which is a unit for reading and writing, and a block, which is composed of several pages and can be erased. Because reading and writing units and erasing units is different, NAND flash memory has the following characteristics [2].

Erase – Before – Write: Flash memory is not overwritable. To write new data onto existing data, you must erase the block containing the data and rewrite it.

Write Amplification: Since the writing of flash memory is performed in accordance with the unit of the page, even a write-operation smaller than the unit of the page causes additional writing to the remaining space. There is a method of avoiding write operation of a unit smaller than a page or storing data in a memory buffer and writing data at a time.

Limit number of records: There is a limit on the number of erasable and writable flash memories due to the limitation of the media material. If the number of erasable blocks exceeds the limit, the corresponding blocks cannot be erased and recorded anymore. In order to prevent this, all data blocks must be written uniformly to each other, and a special algorithm is applied to this, which is called 'wear-leveling'.

2.2 FTL (Flash Translation Layer)

The Flash Translation Layer (FTL) is a system software that hides the internal characteristics of the flash memory and maps Logical Block Address (LBA) and Physical Block Address (PBA) to the user in order to efficiently use a memory interface that is not suitable for flash memory [2].

2.3 FAST FTL (Fully Associative Sector Translation FTL)

FAST FTL is an FTL using Log Block Scheme which is a method to temporarily store updated data in a log block. Since all data blocks can be mapped to one log block, it shows good performance for write requests to random addresses [3].

FAST FTL consists of Block Mapping Table using Hybrid Mapping Scheme and Log Mapping Table which is mapping information of Log Block. The Log Mapping Table consists of the original address of the data, the address of the updated data, and the valid mark.

When searching for the latest data in the Log Mapping Table, it contrasts with the LPA of the data to be read in the Log Mapping Table. If the LPA exists in the Log Mapping Table, the updated data is read. If not, the data is read from the original address.

3 Indexing Log Areas

3.1 Structure of Log Mapping Information Indexing

In suggested idea, Log Mapping Table is indexed by the BST structure based on LPN. If the LPN is searched for and checked in the index, it is possible to retrieve the updated page data without searching for additional log mapping information based on the PBN and page location.

3.2 Indexing Process

In the indexing process, the allocation time of the log block is firstly compared with the timestamp, and the search is started from the most recently allocated log block.

Algorithm of creating log block index

```
Input: Log Mapping Block (LMB)
Output: BST
Procedure: Log_Indexing (LMB)
BST ← Binary Search Tree consist of (LPA, PBN, offset)
for PBN: physical block number of log blocks
  for LPA, offset, valid: LPA, valid mark in log blocks
    if valid mark is valid then
    BST.add(LPA, PBN, offset);
    end if
  end for
end for
Return BST;
```

According to this algorithm, after generating the BST based on the LPN, the PBNs stored in the log mapping information are sorted in reverse order of the generated time, and if the valid mark of the corresponding LPA is valid, the PBN is added to the BST.

3.3 Deciding Whether Data Is Updated Using the Index

By implementing the index according to the above procedure, you can assign the LPN that you want to search to the index to find out whether it is up-to-date and where the latest data exists. This procedure follows the procedure of under algorithm.

Algorithm of discover whether to update using index

```
Input: Logical Page Number discovering whether to update (LPN)
Output: PBN, offset
Procedure: Search_Index (LPN)
BST ← Binary Search Tree consist of (LPA, PBN, offset)
while BST is not null
  if LPN equals BST.LPN then
    Return BST.PBN, BST.offset
  else
    BST ← BST.next;
  end if
end while
Return get data from mapping table;
```

4 Experimental Results

In the experimental environment, FAST FTL simulator was implemented to measure the read performance of FAST FTL. [5] The flash memory size is 10.3 MB, the total number of blocks is 640, the number of pages per block is 32, and the size of one sector is 528 bytes. This experiment is a method of indexing log mapping information, and the number of read, write, and erase operations of FTL is substantially the same. Therefore, the number of times the LPN of the log mapping information is compared with the LPN to be searched is regarded as a performance index.

4.1 Comparison by Log Block Ratio

When the number of data blocks is 500, and the number of log blocks is 5, 10, 15, 20, and 25, respectively, the number of LPN comparison counts of non-indexed search and indexed search is measured. The LPN in the Log Block was randomly updated and all data was read sequentially (Table 1).

Table 1. Number of LPN comparisons based on log block ratio

Log/data	Common method	Proposed method
1% (5/500)	2,547,283	151,180
2% (10/500)	5,069,247	169,317
3% (15/500)	7,565,691	174,210
4% (20/500)	10,038,861	187,388
5% (25/500)	12,510,098	210,923

The average time complexity of the Binary Search Tree is O(log n) and the total number of logs in the FTL is (number of Log × Page per Block) called NL in this paper. And the average time complexity of the non-indexed search is O(NL), and time complexity of the search using the index is O(log NL). Therefore, the number of comparisons increases linearly as the ratio of the general search increases, and that the index increases as the logarithmic scale.

4.2 Comparison Based on the Weight of Frequently Updated Data

In FAST FTL, data that is frequently updated is called Hot Data, and data that is not updated frequently is called Cold Data. The number of data blocks is 500, the number of log blocks is 5, and all the data are sequentially read. Table 2 is the result of measuring the number of LPN comparison according to the proportion of hot data. The number of comparisons increases linearly as the ratio of the general search increases, and that the index search increases as the logarithmic scale.

Table 2. Number of LPN comparison based on the weight of frequently updated data

Hot data/all data	Common method	Proposed method
100%	2,547,283	151,180
50%	2,547,547	131,363
20%	2,547,889	96,354
10%	2,547,687	86,901
1%	2,550,926	80,447

In the normal FAST FTL, the log area is searched regardless of the ratio of Hot Data and Cold Data, so it is not influenced by the number of comparisons. Since the depth of nodes in the binary search tree is closely related to the rate of hot data, the number of LPN comparisons decreases as the weight of hot data decreases.

5 Conclusion

As the utilization rate of the flash memory increases, the capacity and the speed are also improved. FAST FTL has been studied to efficiently use the interface between flash memory and existing file system. It has been necessary to search the log area to see if the data is up to date. In this paper, we propose a method to index the log area in order to avoid searching the entire log area when it is updated. We compare the proposed method with the ratios of the log block and the frequency of data usage and show the validity of the proposed method.

The ratio of log data defined in the proposed method is based on the log data ratio of the flash memory currently used and the ratio of hot data is examined from 1% to 100% considering various situations. When index is used, the number of LPN comparisons increases in the form of logarithmic scale, which is more efficient than the search algorithm that increases to the conventional linear form. It means that the flash memory technology that has been developed so far can easily cope with the increase in the size of the log area to be searched as the capacity of the flash memory increases. Also, the fact that the number of LPN comparisons is reduced according to the ratio of hot data is easier than the general FAST FTL when data must be stored in flash memory for a long time. From the above point of view, it will contribute to the maintenance of search performance according to the development of flash memory in the future.

Acknowledgement. This research was supported by the MIST (Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP for Information & communications Technology Planning & Evaluation) (2015-0-00908).

References

1. Kawaguchi, A., Nishioka, S., Motoda, H.: A flash-memory based file system. In: USENIX, pp. 155–164 (1995)
2. Intel Corporation: Understanding the Flash Translation Layer (FTL) specification. Application Note AP-684 (1998)
3. Lee, S.W., Park, D.J., Chung, T.S., Lee, D.H., Park, S., Song, H.J.: A log buffer-based flash translation layer using fully-associative sector translation. ACM Trans. Embed. Comput. Syst. **6**, 18 (2007)
4. Kang, W.H.: On fast searching of log area in FAST FTL using counting bloom filters. M.S. thesis. SungKyunKuan University, Seoul (2010)
5. YuSangHyeok/fast-ftl (Github). http://www.github.ocm/YuSangHyeok/fast-ftl

Aspects of Intercultural Communication in IT: Convergence of Communication and Computing in the Global World of Interconnectedness

Marcel Pikhart(✉)

Faculty of Informatics and Management, University of Hradec Kralove, Rokitanskeho, Hradec Kralove, Czech Republic
marcel.pikhart@uhk.cz

Abstract. The paper focuses on the influence of culture and communication on the way IT processes are managed. The global world is very much intercultural and interconnected, mostly due to the vast impact of the Internet communication. To communicate efficiently, the intercultural aspects of information transferred by computerized information systems must be taken into account. There are many issues which have to be addressed by the IT departments in companies if they want to be successful and competitive. The paper stresses the importance of interculturality in the global IT world and tries to find key issues which are crucial to stability of IT processes in companies. There are also very important issues connected to computational linguistics and machine learning which can be utilized when developing tools for intercultural global world because both computational linguistics and machine learning can bring new ideas and issues to areas such as eLearning. In eLearning processes the use of artificial intelligence though computational linguistics and machine learning can eliminate the potential bottlenecks which we are now facing globally.

Keywords: Business communication · Technology and communication · Global communication · Communication studies · Information management · Pragmalinguistics · Computational linguistics

1 Introduction

The continuously evolving environment of IT brings many challenges and also threats in the global world which might still be neglected by IT departments in companies [1, 4, 6, 7]. Interculturality as the global phenomenon [2, 3, 15, 16] brings many opportunities but also threats to the IT industry, and we have to address them appropriately to eliminate potential drawbacks which could lead to inefficiency and slowed down processes which will not be functional effectively. In the past few years, IT has influenced and changed every aspect of our lives, however, the unprecedented changes in the world around us have also influenced IT [5, 8, 9]. The people and cultures are more and more connected and this can be described as interculturality and interconnectedness.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 251–256, 2020.
https://doi.org/10.1007/978-981-32-9244-4_35

This new phenomenon needs our attention and should be reflected by IT professionals all over the word [9, 11, 14]. Intercultural skills are, therefore, needed not only by people who communicate personally with other people from other cultures, but in a large scale by people who study IT and who create IT strategy in our companies. More and more people come into contact with people from other cultures, technology makes it possible to travel long distances, the Internet connects people across cultural, national, historical and economic boundaries, people travel much more than ever before and workforce is becoming extremely diverse [1, 2, 10]. All these factors massively influence the way people communicate by using electronic devices and how they share data over the Internet, which has become ubiquitous. The fact that we live in the global village makes our communication very fast but also vulnerable due to the massive diversity in the communication channels and the speed flow of information [12].

Without the knowledge and awareness of these modern issues, IT departments and IT specialists cannot succeed, and thus, the companies will be facing dramatic threats which will inevitably lead to the loss of business opportunities [13]. The author of the paper works as both an academic and an independent consultant and the paper, therefore, tries to connect theoretical viewpoints and their practical implications.

The aforementioned interculturality is manifested in IT sector quite clearly. IT is a very important tool for marketing, internal communication, PR communication, etc., and all these aspects of business are in today's global world influenced by diversity and interculturality [9, 11]. E.g., advertisers have to take into account cultural differences when planning advertising campaigns, have to consider physical layout of the text on the website, the appropriate use of visuals, the use of colours, etc., because all these aspects and items will have significantly different meaning for example in Europe and in Asia. This is the most important manifestation of interculturality – i.e. different understanding of symbols and visual messages used massively in IT industry, be it websites, advertisement, internal communication tools, etc. [10].

There are more practical issues for IT such as:

- Respect taboos and customs,
- Avoid references that will not be understood by people from other cultures,
- Check that the symbols or icons are understood in the target market in the way you intend, and
- Consider the balance between text and visual – this will need to be changed when you move into other culture, etc. [1, 2].

All these, and many more phenomena, must be realized and then carefully considered by IT professionals if they want to create useful and powerful tools for companies and subsequently enhance their efficiency in the current global information environment.

2 Research

The initial idea of the presented paper was to test the awareness of IT departments in several multinational corporations doing business in Europe and compare the results with similar companies in Asia (China and Taiwan). Subsequently, it was realized that

there was a need for another research into the awareness of the students, therefore, similar research was conducted in European (namely Czech) university students of IT and Taiwanese university students of IT. The comparative research will give us the opportunity to see the reality and the potential problems which are connected to the researched area.

The aim of the research was to map the situation and flag up potential bottlenecks which could cause inefficiency in the IT departments not directly, but It departments creating websites and apps which will not be efficient. It should also show if there are any regional differences (Europe versus Asia) and generation differences (students of IT versus IT employees, which are usually two different generations).

This kind of research is unique as there is no comparative research into the topic of interculturality in IT and the awareness of the topic is limited. A surprising fact is that in business sector intercultural business communication has been present, researched and taught for more than three decades, however, the issue in IT still lacks any attention, despite its urgency in the global world. IT professionals are responsible for web creation, apps, global marketing etc. and they should definitely be aware of the potential problems.

2.1 Research Description

The qualitative research through interviews was conducted in several multinational European companies (total number 7) and also several (3) multinational companies doing business in China and Taiwan into the awareness of intercultural issues of their IT employees. The same research was also conducted in the group of university IT students, both in the Czech Republic and in Taiwan. The research was conducted during the years 2017 and 2018. The results were collected by means of an online questionnaire. Standard data processing methodology was used, data analysed, compared and findings created.

2.2 Research Hypothesis

The hypothesis was that the level of awareness of the IT employees and IT students of the importance of interculturality will not be dramatically different in Europe and China due to very similar educational systems in these two regions.

2.3 Research Results

The results of the research were somehow appalling. Whereas in Europe the intercultural issues were clearly realized by both IT employees and university students of IT, however, in Asia the awareness was really limited.

The research clearly showed that the majority (86%) of European companies, IT employees and also IT students are clearly aware that people communicate very differently when they come from different cultures and also proved to know that all these phenomena are very much manifested and reflected in IT area, however, in Asia this awareness is radically limited (only a few respondents (23%) knew that interculturality could influence the transfer of information significantly).

The vast majority of European respondents (90%) showed their interest in the topic and also stated that they had been trained on intercultural issues in IT (67%) or at least they knew that these aspects are very important, whereas Asian respondents did not realize the importance.

All these findings lead us to the idea that it is necessary to improve our educational system which still lacks sufficient support for the students of IT to equip them with relevant expertise which will help them to survive in the global environment which is very different from the environment twenty years ago. IT department have to realize the importance of the issue of interculturality and if they want to create tools which may be efficient and successful, they have to bear these phenomena in their mind.

2.4 Research Limitations

The research was conducted in two regions (Central Europe and Asia) and despite the fact of this geographical diversity, other regions may show different results. However, it can be claimed, that the results are transferrable into the global situation and therefore useful and applicable in larger scale. There is also a lot of space for further research into intercultural issues in IT, and the author of this paper himself focuses on further and more detailed study of the current situation with the prospect of new findings which could be applied in university IT curricula and IT industry.

3 Discussion

The research clearly showed that there are significant differences in the given regions. Europe proving to be well informed, whereas Asia not very much aware of the importance of the issue. The paper attempts to show that intercultural issues in IT need our undivided attention if we want to be competitive, because various cultural clashes through IT environment are very sensitive and can easily harm the company reputation.

When creating websites, preparing eLeaning courses, creating apps, or any platform which will be used globally, which is nowadays basically any product of IT, computational linguistics and machine learning could be used to generate tools which will be culturally and interculturally neutral or the given audience. Artificial intelligence and its implementation in all these platforms will also be able to accommodate to the given culture so that the app or platform will be neutral and will not do any harm, mostly in culturally sensitive areas such as China.

China is rising dramatically in the global market such as the USA in the 1980s and if it wants to succeed as a global superpower, and if it wants to be successful as a trustworthy partner for American and European companies, it must prove that it is able to accommodate to the needs of the global world. When the USA started to do business in the 1980s with China, many courses of intercultural communication were introduced in the US universities because the companies realized that if they wanted to be successful in the given region, they must find appropriate ways of communication there. Similarly, IT services and their professionality is the key to financial gains and global success.

4 Conclusion

IT is ubiquitous in the global market and the whole IT industry must realize that its professionality is the key to global competitiveness and when we talk about global issues they are nowadays always connected to diversity and interculturality. Therefore, understanding and implementing critical communication strategies to IT sector will necessarily enhance global chances of competitiveness and profitability. The companies realized the need for this awareness in business already thirty years ago when the world became more interconnected, however, the IT sector is still lagging behind and waiting. Therefore, this paper attempts to highlight the fact that being competitive and successful globally is only through the respect for the differences and diversity. IT sector is more than any other sector operating globally, therefore, this is a very urgent task for it more than for any other business area.

Acknowledgments. The paper was created as a part of SPEV 2019 project at the Faculty of Informatics and Management, University of Hradec Kralove, Czech Republic. The author thanks Jan Sprinar for his cooperation.

References

1. Cerna, M., Svobodova, L.: Internet and social networks as the support for communication in the business environment - pilot study. Hradec Econ. Days **7**(1), 120–126 (2017). Edited by Jedlicka, P., Maresova, P., Soukal, I.
2. Eckert, S.: Intercultural Communication. Thomson, Mason (2006)
3. Gibson, R.: Intercultural Business Communication. Oxford University Press, Oxford (2002)
4. Klimova, B.: Smart teacher. In: Smart Education and E-Learning 2017, SEEL 2017. SIST, vol. 75, pp. 321–328 (2018)
5. Klimova, B., Poulova, P.: Mobile learning and its potential for engineering education. In: IEEE Global Engineering Education Conference, EDUCON, April 2015, Article no. 7095949, pp. 47–51 (2015)
6. Klimova, B., Poulova, P.: Older people and technology acceptance. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). LNCS, vol. 10926, pp. 85–94 (2018)
7. Maresova, P., Klimova, B.: Economic and technological aspects of business intelligence in european business sector. In: 11th International Scientific Conference on Future Information Technology (FutureTech)/10th International Conference on Multimedia and Ubiquitous Engineering (MUE). Advanced Multimedia and Ubiquitous Engineering: Futuretech & MUE. Lecture Notes in Electrical Engineering, Beijing, Peoples Republic of China, 20–22 April 2016, vol. 393, pp. 79–84 (2016)
8. Mohelska, H., Sokolova, M.: Smart, connected products change a company's business strategy orientation. Appl. Econ. **48**(47), 4502–4509 (2016)
9. Pikhart, M.: Sustainable communication strategies for business communication. In: Soliman, K.S. (ed.) Proceedings of the 32nd International Business Information Management Association Conference (IBIMA), Seville, Spain, 15–16 November 2018, pp. 528–553. International Business Information Management Association (2018)

10. Pikhart, M.: Intercultural business communication courses in european universities as a way to enhance competitiveness. In: Soliman, K.S. (ed.) Proceedings of the 32nd International Business Information Management Association Conference (IBIMA), Seville, Spain, 15–16 November 2018, pp. 524-527. International Business Information Management Association (2018)
11. Pikhart, M.: Multilingual and intercultural competence for ICT: accessing and assessing electronic information in the global world. In: MISSI 2018. Advances in Intelligent Systems and Computing, vol. 833, 2019, pp. 273–278 (2018)
12. Pikhart, M.: Technology enhanced learning experience in intercultural business communication course: a case study. In: Hao, T., Ha, W.C. (eds.) Emerging Technologies for Education. Third International Symposium, SETE 2018, Held in Conjunction with ICWL 2018, Revised Selected Papers. Lecture Notes in Computer Science. Springer (2018)
13. Pikhart, M.: Communication based models of information transfer in modern management - the use of mobile technologies in company communication. In: Soliman, K.S. (ed.) Proceedings of the 31st International Business Information Management Association Conference (IBIMA), Milan, 25–26 April 2018, pp. 447–450. International Business Information Management Association (IBIMA) (2018). ISBN 978-0-9998551-0-2
14. Pikhart, M.: Managerial communication and its changes in the global intercultural business world. In: Special Issue on Education. SHS Web of Conferences, ERPA 2017, vol. 37 (2017). https://doi.org/10.1051/shsconf/20173700001
15. Samovar, L., Porter, R., McDaniel, E.R.: Communication Between Cultures. Thomson, Belmont (2012)
16. Washington, M.C., et al.: Intercultural communication in global business: an analysis of benefits and challenges. Int. Bus. Econ. Res. J. **11**(2), 217–222 (2012)

Improvement of QSL by Ontologies of E-Questionnaire, E-Testing, and E-Voting Systems

Yuan Zhou, Yuichi Goto(✉), and Jingde Cheng

Department of Information and Computer Sciences, Saitama University, Saitama 3388570, Japan
{shuugen, gotoh, cheng}@aise.ics.saitama-u.ac.jp

Abstract. QSL is the first specification language for specifying various e-questionnaire, e-testing, and e-voting systems. Although the terminologies among systems and services of e-questionnaire, e-testing, and e-voting are different, QSL has explicitly summed up three kinds of terminologies into one, so that causes QSL has poor usability. The ontologies to summarize the terminologies in e-questionnaire, e-testing, and e-voting systems, to find out the corresponding relations with terminology of QSL, and to clarify the relations of e-questionnaire, e-testing, and e-voting systems can improve usability of QSL, so that the stakeholders can communicate and write requirement specifications easily. However, there is no ontology of e-questionnaire, e-testing, and e-voting systems. This paper proposes the ontologies for e-questionnaire, e-testing, and e-voting systems. Based on the ontologies, we present the improvement of QSL, so that stakeholders can use arbitrary specific terminologies to specify the requirement specifications for other kinds of systems and services.

Keywords: Ontology · E-Questionnaire · E-Testing · E-Voting · Specification language

1 Introduction

Questionnaire, testing, and voting are the essential activities of the modern communities as the general and indispensable methods for a group of people to express a choice, or to assess people's ability [5, 8]. Over two decades, many kinds of questionnaires, testing, and voting are performed in some completely electronic ways to do questions and answers. Electronic questionnaire, electronic testing, and electronic voting (e-questionnaire, e-testing, and e-voting for short) are indispensable electronic services in our society. A huge variety of e-questionnaire, e-testing, and e-voting systems has been designed, developed, maintained, and operated in ad-hoc ways. If these systems are unreliable, lower security, strange in use, it will have a serious impact on our society. There is still an important research topic of how to design, develop, maintain, and operate reliable, highly secure, and user-friendly e-questionnaire, e-testing, and e-voting systems.

In addition, because e-questionnaire, e-testing, and e-voting have common processes, that is, from preparing questions, following by authenticating respondents,

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 257–264, 2020.
https://doi.org/10.1007/978-981-32-9244-4_36

through submitting answers, and ending to analyzing, counting, and declaring results, the systems that provide those services have common functions to do the processes. In fact, some representative systems [1, 6, 7, 10, 11, 13, 15–17] exist to provide three-in-one service for people all over the world.

The mutual collaboration of the stakeholders is the foundation for the development and operation of e-questionnaire, e-testing, and e-voting systems. However, there is no communication tool shared among the stakeholders of systems and services of e-questionnaire, e-testing, and e-voting. As a result, the stakeholders are difficult to communicate to implement the systems, because there is neither an exhaustive requirement list to have a grasp of the overall nor a standardized terminology for those systems and services to avoid ambiguity.

QSL [20] is the first specification language as a communication tool for specifying various e-questionnaire, e-testing, and e-voting systems with a standardized, consistent, and exhaustive list of requirements so that the stakeholders can communicate easily and unambiguously, and deal with and describe the requirement specifications for three kinds of systems and services. However, the terminology of QSL is designed according to e-questionnaire systems, and extended based on e-testing and e-voting systems. As a result, the stakeholders among e-testing and e-voting systems are difficult and not willing to use QSL. Thus, QSL has poor usability.

To improve QSL about its usability so that the stakeholders can communicate and write requirement specifications by familiar terminology, the solution is to define the ontologies to make clear corresponding relationships of systems and services of e-questionnaire, e-testing, and e-voting. However, there is no ontology for those systems and services.

In this paper, we proposed ontologies that are from collecting the terminologies of representative e-questionnaire, e-testing, and e-voting systems, through organizing the corresponding relationships of these three kinds of systems and services. Based on the ontologies, we presented the improvement of QSL by a reference list of replacements for terminologies, so that stakeholders can use arbitrary specific terminology to specify the specifications for other e-questionnaire, e-testing, and e-voting systems. Section 2 introduces QSL and its terminology. Section 3 shows the ontologies of e-questionnaire, e-testing, and e-voting systems. Section 4 presents the improvement of QSL. Some conclusions and future works are shown in Sect. 5.

2 Questionnaire Specification Language (QSL)

2.1 Overview of QSL

Questionnaire Specification Language (QSL) serves as a formalized specification for specifying various e-questionnaire, e-testing, and e-voting systems [20–23]. QSL is based on Extensible Markup Language (XML) [18]. We have evaluated QSL about description power to ensure its completeness manifesting in specifying various e-questionnaire, e-testing, and e-voting systems. Current QSL can cover more than 95% system requirements and 95.4% service requirements, and provides enough

notations to describe the requirements for data portability, and can cover the existing specifications compared with the related works [24].

The tags for QSL are used to provide terminology to describe requirements of e-questionnaire, e-testing, and e-voting systems. Using the tags, QSL schemas are used to constrain the requirements in a formal way. QSL schema is a collection of requirements formalized by XML schema, and clear definition of the relationship among the requirements. The schema through the XML parser, the user can get the QSL templates that are the requirements formalized by XML, and the requirements correlate with corresponding necessary requirements. The users choose the desirable QSL schema, through XML parser, they can get the QSL template, and then they input appropriate values can get formalized requirement specifications.

QSL can be used in three ways. Firstly, QSL can be used to specify e-questionnaire, e-testing, and e-voting systems. In other words, QSL can be used to specify the system requirements of those systems. Secondly, QSL can be used to specify e-questionnaire, e-testing, and e-voting on the system, i.e., QSL can be used to describe requirements related to each phase of these services, that is, from preparing questions, following by authenticating respondents, through submitting answers, and ending to analyzing, counting, and declaring results. Thirdly, QSL can be used as format for portable data of e-questionnaire [9], e-testing, and e-voting. Formats of data of e-questionnaire, e-testing, and e-voting systems are the parts of specifications of e-questionnaire, e-testing, and e-voting systems. Portable data [4] is formatted according to a published syntax and where the metadata is explicit, either included with the data or by reference to an open technical dictionary.

2.2 Current Terminology of QSL

Current QSL (version 3.1) provides 93 tags as terminology [2]. There are three kinds of tags, which are used to specify common requirements and specific requirements, and to construct the document.

Current terminology for QSL is defined according to e-questionnaire systems, and extended based on e-testing and e-voting systems. Because e-questionnaire, e-testing, and e-voting systems have lots in common, most of the tags are defined according to the terms of e-questionnaire and e-questionnaire systems as the common tags. The rest of a few tags are extended and defined according to the terms of systems and services of e-testing and e-voting. However, the stakeholders who use QSL are possible to confuse, after all, they are not proficient in e-questionnaire, e-testing, or e-voting. For example, in QSL, we defined a tag named *questioner* that is a participant role who designs a paper sheet and settings. It is normally called *examiner* in e-testing and the stakeholders in e-testing prefer to and are familiar with examiner. Thus, even though we provide the exhaustive requirement list and terminology for the stakeholders, the current terminology of QSL is not going to raise the efficient communications among the stakeholders.

3 Ontologies of E-Questionnaire, E-Testing, and E-Voting Systems

3.1 Investigation of Terminologies

In recent years the development of ontologies has been moving from the realm of AI laboratories to the desktops of domain experts. Ontology is an explicit specification of a conceptualization, and it defines terminologies for researchers who need to share information in a domain and includes machine-interpretable definitions of basic concepts in the domain and relations among them [12].

In order to improve the usability of QSL, we propose the ontologies to clearly define the corresponding relationships of systems and services of e-questionnaire, e-testing, and e-voting. We investigated 26 e-questionnaire, 23 e-testing, and 25 e-voting systems [20], which are representative systems seizing a large number of high quality customers all over the world for serving a relatively long time. We also enumerated and summarized the terms. There are 179 terms for e-questionnaire, e-testing, and e-voting systems. Each term has its relationship with others. We summarized and organized the corresponding relationships in different fields of e-questionnaire, e-testing, and e-voting. We defined ontologies for e-questionnaire, e-testing, and e-voting systems, respectively.

3.2 Ontologies

The ontologies are expressed in the Web Ontology Language (OWL) [19] for describing relationships for systems and services of e-questionnaire, e-testing, and e-voting. OWL is an XML-based language for publishing and sharing ontologies. We used Protégé [14] to develop the ontologies.

A simple ontology for classes in the top level and their relationships is depicted in Fig. 1. It is a main, general, and common part for the ontologies for e-questionnaire, e-testing, and e-voting systems. All the common classes for e-questionnaire, e-testing, and e-voting are constructed according to this figure. The system contains environment, function, security, and data. The service consists of paper sheet, settings, and data. The relationships among the classes are defined as object properties. Besides, e-voting does not need a logic class.

Figure 1 illustrates the common classes and their relationships of e-questionnaire, e-testing, and e-voting systems. As to the different classes and relationships for three kinds of systems and service in details, we take an example about participant class, which is also illustrated in previous figure. Figure 2 illustrates its subclasses and the different terms in e-questionnaire, e-testing, and e-voting systems, respectively. Some roles only exist in a field, such as “assessor” only exists in e-testing. All the roles are the subclasses of participant, and they have relationships with other classes, such as *Name*, *Affiliation*, etc. In addition, all the roles have relationship with *Authentication*, and have different methods during different phases. The detailed information about the ontologies refers to [3].

Using the ontologies, the stakeholders can directly use their familiar terminologies in different fields of e-questionnaire, e-testing, or e-voting, so that they can easily and

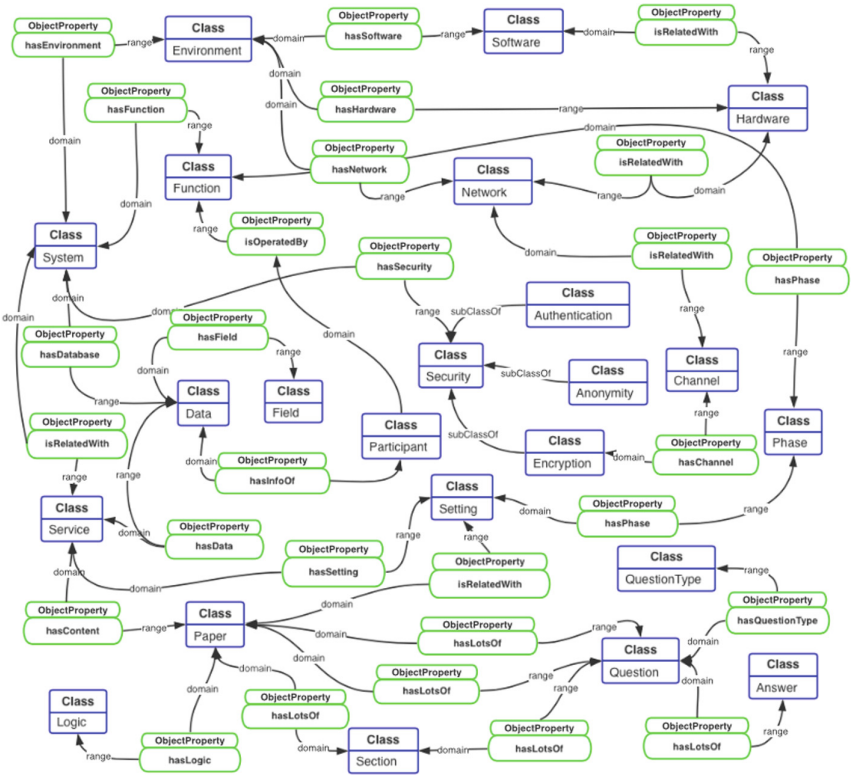


Fig. 1. A figure about a main, general, and common part of the ontologies for e-questionnaire, e-testing, and e-voting systems.

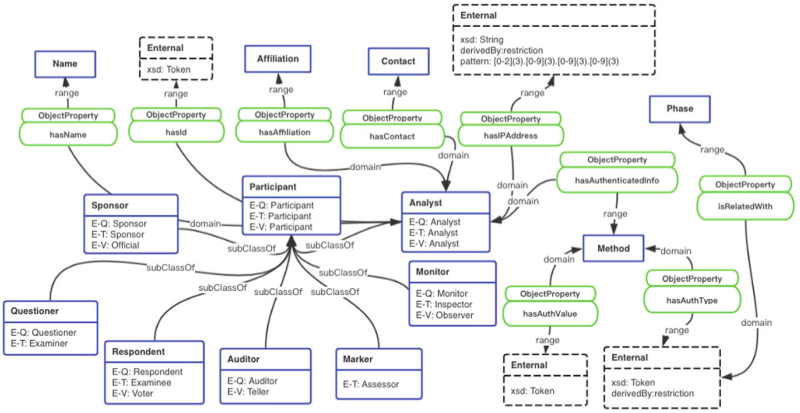


Fig. 2. A figure about the ontologies for participant in e-questionnaire, e-testing, and e-voting systems.

efficiently communicate with each other. Furthermore, from the used terms, the ontologies are used to provide corresponding terminologies for the stakeholder who is judged in which field. For instance, when they uses the term "assessor" to define the participant who marks and gives the scores, the terminology about field in e-testing is provided and replaced in specification. In other words, in common sense, we can judge that this stakeholder is familiar with and prefer to communicate by the terminology of e-testing.

4 Improvement of QSL

Based on the ontologies, we improved QSL. The ontologies are a part of or an attribute of QSL. Using the ontologies can improve the usability of QSL. It helps to indicate the relationships of e-questionnaire, e-testing, and e-voting system so that the stakeholders can share common understandings of the structure of information in these three kinds of systems and services.

We defined a reference table to list the replacements of terminologies of e-questionnaire, e-testing, and e-voting system. An example about a replacement of participant is listed in Table 1. The values listed as none do not need to specify in requirement specification. There are much tags should be replaced, and complete replacement list is represented in [3].

Table 1. A reference list of replacement of e-questionnaire, e-testing, and e-voting systems.

Tags in QSL	Classes in E-Questionnaire	Classes in E-Testing	Classes in E-Voting
Participant	Participant	Participant	Participant
Sponsor	Sponsor	Sponsor	Official
Questioner	Questioner	Examiner	(None)
Respondent	Respondent	Examinee	Voter
Analyst	Analyst	Analyst	Analyst
Monitor	Monitor	Inspector	Observer
Marker	(None)	Assessor	(None)
Auditor	(None)	(None)	Teller

A better solution for improving QSL is showed in Fig. 3 that summarizes the relationship between QSL documents and ontologies, and also shows the usages of them. The stakeholders create content in a QSL template without values. That template conforms to the rules of the QSL schemas. They then use ontology which clearly defines the familiar terminology and can be used to be instead of QSL original terminology.

Considering that the users can easily create a QSL format requirement specification and it is not necessary to know structure and terminology of QSL, we propose to develop a QSL structure editor. Hundreds of requirements and their relationships and constraints for e-questionnaire, e-testing, and e-voting systems have to be taken into

careful consideration. A QSL structure editor to edit and generate QSL format requirement specification should satisfy the following requirements.

R1: The editor must provide services to guide the users to choose the corresponding requirement templates according with their suitable identities.
R2: The editor must provide services to perform the requirement list with a graphical user interface that hide the code in the background and present the content to the user in more user friendly forms and show guidance according to the suitable terminology.
R3: The editor must provide services to guide to fill the appropriate values, and to validate and verify whether the filled values are valid or not in conformance with QSL schemas, and point out the invalid values and defined formats.
R4: The editor must provide services to import the QSL format requirement specification, and display the requirement list in graphical user interface.
R5: The editor must provide services to hide or show the requirements, which are logically associated with a chosen requirement.
R6: The editor must provide services to output QSL format requirement specification replaced by the familiar terminology.

According to the QSL templates, QSL schemas, and the ontologies, the QSL structure editor helps the users to easily write QSL-format requirement specification. Through the QSL structure editor, the QSL template defined by a series of familiar terminology is filled with values that output a requirement specification.

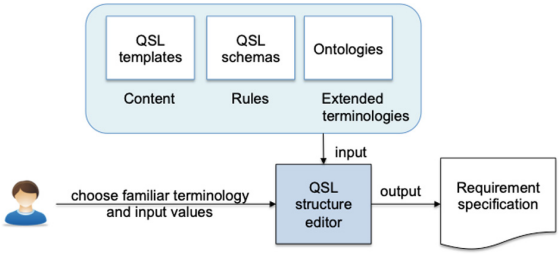


Fig. 3. A figure about the relationship and usage of ontologies in QSL.

5 Concluding Remarks

We have proposed and developed ontologies of e-questionnaire, e-testing, and e-voting systems that clearly defines the relationships of e-questionnaire, e-testing, and e-voting systems. In addition, based on the ontologies, we improved QSL so that stakeholders can use their familiar and arbitrary specific terminologies to specify the specifications for other kinds of systems and services of e-questionnaire, e-testing, and e-voting. In the future, we are continuing improving QSL for investigating much more representative systems for extending the ontologies. In addition, we will implement a QSL structure editor.

References

1. AddPoll. http://www.addpoll.com
2. AISE Lab: QSL manual (version 3.1). http://www.aise.ics.saitama-u.ac.jp/QSL/
3. AISE Lab: Ontologies for e-questionnaire, e-testing, and e-voting systems. http://www.aise.ics.saitama-u.ac.jp/QSL/ontology
4. Benson, P.R.: Data portability: it is about the data and the quality of the data. In: MIT Information Quality Industry Symposium on Proceedings, pp. 614–618 (2009)
5. Cheng, J., Goto, Y., Koide, M., Nagahama, K., Someya, M., Utsumi, Y., Shionoiri, A.: ENQUETE-BAISE: a general-purpose e-questionnaire server for ubiquitous questionnaire. In: 2nd IEEE Asia-Pacific Services Computing Conference on Proceedings, Tsukuba, Japan, pp. 187–194. IEEE Computer Society Press (2007)
6. Constant Contact. http://www.constantcontact.com
7. eSurv. http://eSurv.org
8. Goto, Y., Cheng, J.: Information assurance, privacy, and security in ubiquitous questionnaire. In: 4th International Conference on Frontier of Computer Science and Technology on Proceedings, Shanghai, China, pp. 619–624. IEEE Computer Society Press (2009)
9. Kamata, Y., Goto, Y.: Improvement of data portability of ENQUETE-BAISE: a general-purpose e-questionnaire server for ubiquitous questionnaire. In: 2018 IEEE Smart World Congress on Proceedings, Guangzhou, China, pp. 174–179. IEEE Computer Society Press (2018)
10. Mobo Survey. http://www.mobosurvey.com
11. MySurveyLab. http://enq-maker.com
12. Ontology development 101: a guide to creating your first ontology. https://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html
13. ProProfs. http://www.proprofs.com
14. Protégé. https://protege.stanford.edu
15. QuestionPro. http://www.questionpro.com
16. So Jump. http://www.sojump.com
17. SurveyMonkey. http://www.surveymonkey.com
18. W3C: Extensible Markup Language (XML) 1.0, 5th edn. http://www.w3.org/TR/2008/REC-xml-20081126/
19. W3C: OWL Web Ontology Language Document Overview, 2nd edn. https://www.w3.org/TR/owl2-overview/
20. Wang, Z., Zhou, Y., Wang, B., Goto, Y., Cheng, J.: An extension of QSL for e-testing and its application in an offline e-testing environment. In: Park, J., Chao, H., Arabnia, H., Yen, N. (eds.) FutureTech 2015. LNEE, vol. 352, pp. 7–14. Springer, Heidelberg (2015)
21. Zhou, Y., Gao, H., Cheng, J.: QSL: a specification language for e-questionnaire, e-testing, and e-voting systems. In: Park, J., Jin, H., Jeong, Y., Khan, M. (eds.) FutureTech 2016. LNEE, vol. 393, pp. 255–261. Springer, Singapore (2016)
22. Zhou, Y., Gao, H., Cheng, J.: An extension of QSL for e-voting Systems. In: Park, J., Pan, Y., Loia, V. (eds.) CSA2016. LNEE, vol. 421, pp. 87–96. Springer, Singapore (2016)
23. Zhou, Y., Goto, Y., Cheng, J.: QSL: a specification language for e-questionnaire systems. In: 5th IEEE International Conference on Software Engineering and Service Science (ICSESS 2014) on Proceedings, Beijing, China, pp. 224–230. IEEE Computer Society Press (2014)
24. Zhou, Y., Matsuura, D., Goto, Y., Cheng, J.: Evaluation about the descriptive power of QSL: a specification language for e-questionnaire, e-testing, and e-voting systems. In: 2018 IEEE SmartWorld International Conference on Proceedings, Guangzhou, China, pp. 198–203. IEEE Computer Society Press (2018)

Exploring eLearning for Dementia Care

Blanka Klimova(✉)

Faculty of Informatics and Management, University of Hradec Kralove,
Rokitanského 62, 500 03 Hradec Kralove, Czech Republic
blanka.klimova@uhk.cz

Abstract. The purpose of this review article is to explore the exploitation of eLearning as caregivers' support for patients with dementia in order to help deliver better care for this target group of people, as well as maintain quality of life of their caregivers. The methods include a method of literature search of available sources describing this issue in the world's databases Web of Science, Scopus, and PubMed, and a method of comparison and evaluation of the findings in the selected studies on the research topic. The findings indicate that the eLearning programs can have a significant impact on the improvement of depressive symptoms of caregivers looking after the patients with dementia. However, more evidence-based studies are needed on this topic.

Keywords: eLearning · Dementia · Caregivers · Benefits

1 Introduction

Nowadays, dementia affects 50 million people worldwide and the forecasts indicate that this number should reach 131.5 million by 2050 [1] (consult Fig. 1 below).

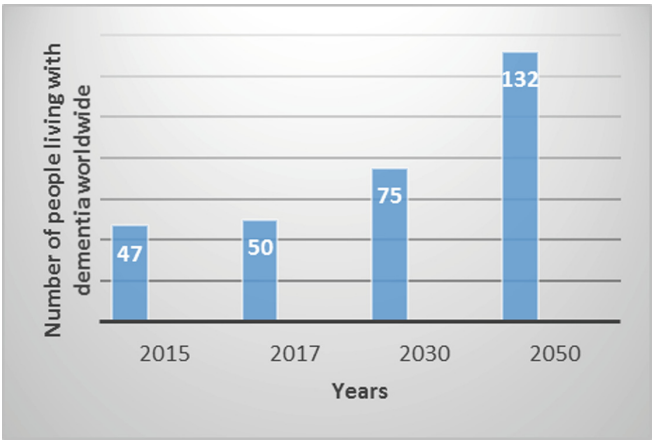


Fig. 1. Number of people living with dementia worldwide (author's own processing, based on [1]).

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 265–270, 2020.
https://doi.org/10.1007/978-981-32-9244-4_37

Dementia is especially characterized by progressive neurodegeneration that results in cognitive decline [2]. Other symptoms of dementia include a considerable loss of memory, orientation problems, impaired communication skills, depression, behavioral changes, and confusion [3, 4]. In addition, people with dementia suffer from behavioral disorders such as apathy, depression, delusions, hallucinations, aggression, irrelevant sexual behavior or sleeping problems.

All these symptoms impose a significant emotional and physical burden not only on the patients themselves, but also on their formal and informal caregivers who have to take care of these patients because they are dependent on their help [5]. At present, there is a tendency to shift the institutional care to the community care [6]. In fact, 80% of the caregivers are family members whose quality of life rapidly declines after taking care of their loved ones suffering from dementia [7]. These informal caregivers usually spend between 5 and 20 h of care per day assisting their family members with dementia [8]. In addition, these caregivers suffer from a gradually increasing physical, mental and economic burden [9]. Therefore, there is an urgent need and effort to alleviate this burden for the informal caregivers and provide them with some professional help. In this respect, technologies seem to be one of the solutions. They have penetrated in all spheres of human life, including learning, which began to be supported electronically in the so-called eLearning.

The research studies emphasize the importance of eLearning for the dissemination of knowledge, understanding particular health issues, continuous education, and training of busy healthcare professionals who wish to access educational programs to maintain or extend their knowledge in response to service needs [10–12]. Most recent trends indicate that eLearning courses also play an important role in inter-professional education and collaboration [13]. The eLearning courses might improve inter-professional care in terms of better communication between healthcare team members, enhanced quality of care, and better outcomes for patients [14].

The purpose of this study is to explore the exploitation of e-learning as caregivers' support for patients with dementia in order to help deliver better care for this target group of people, as well as maintain quality of life of their caregivers.

2 Methods

The methodology of this study follows Moher et al. [15]. The topic of eLearning and its exploitation as a valuable caregivers' support for patients with dementia was searched in the world's databases Web of Science, Scopus, and PubMed in the period of 2010 till 2017 when majority of the studies originated. However, for the comparative reasons, older research studies were used in the Introductory and Discussion part. Furthermore, other relevant studies were reviewed on the basis of the reference lists of the research articles from the searched databases. The key search words were *eLearning* AND *caregivers, eLearning* AND *dementia*. Altogether 82 research studies were identified in the databases mentioned above. Most of them were found in PubMed. After a thorough review of the titles and abstracts and the duplication of the selected studies, 17 studies remained for the full-text analysis, and eventually, five met the set criteria.

The detailed analysis included only those studies, which involved patients with dementia and their caregivers; focused on the use of eLearning as a support tool for caregivers; and were written in English. The studies, which focused on a different target group and disease, were excluded. The same was true for the studies, which were not proper articles, but, for example, descriptive studies discussing the eLearning course for dementia caregivers.

3 Results

Altogether five studies dealing with the research topic were detected. Two studies were randomized controlled studies [16, 17], two were pilot studies [18, 19], and one study was a longitudinal cohort study [20]. The main topic in most of them concentrated on the use of eLearning in the improvement of psychological and emotional distress when caring about people with dementia, as well as dissemination of knowledge about the disease itself and its management. All of them used standardized methods of assessment such as pre and post-tests, online questionnaires, and statistical analysis. Apart from the thousands of participants in the cohort study by Moniz-Cook et al. [20], the detected studies had relatively small samples of subjects, ranging from 35 to 279. The same was also true for the period of assessment, which usually lasted 3-4 months without the follow-up period, except again for [20]. As far as the country of the origin is concerned, four studies were of mixed European origin (Denmark, the Netherlands, Germany, Sweden, UK, Spain, Poland) and one was Chinese study. Overall, participants were satisfied with the e-learning platform, program, or the course.

4 Discussion

The findings from the detected studies show that eLearning programs/courses can have a significant impact on the improvement of depressive symptoms and maintenance of feelings of empathy, understandings, and concern among informal caregivers [16–20]. This is in line with other review studies on this topic [8, 21, 22]. By using eLearning, for example, the virtual reality intervention Through the D'mentia Lens (TDL) can experience what dementia is like and how people with dementia feel [19].

Moreover, eLearning programs/courses can enhance caregivers' knowledge and skills about dementia and its management (cf. [16–18]). As Kurz et al. [23] state, the eLearning, respectively the Internet-based interventions, are a novel format of delivering health-related education, counselling, and support. Research (cf. [18]) indicates that such eLearning programs or courses contribute to the increase of caregivers' confidence, which consequently results in the alleviation of their distress [19]. In addition, by using video cases, the user can practice the best way of communicating [24]. Currently, there is a number of ongoing initiatives, which deal with the eLearning programs such as the RHAPSODY (Research to Assess Policies and Strategies for Dementia in the Young) project, which aims to provide an internet-based information and skill-building program for family caregivers. The eLearning program focuses on managing problem behaviors,

dealing with role change, obtaining support, and looking after oneself [23].Another project has been developed by Moreno et al. [25]. It proposes the design and implementation of a cloud-inspired personalized and adaptive platform, which would offer support to informal caregivers of older people with cognitive impairment. Through a holistic approach comprising technologies and services, which would address the intelligent and interactive monitoring of activities, knowledge management for personalized guidance and orientation, virtual interaction, eLearning, care coordination facilities and social network services, iCarer focuses on the reduction of the informal caregivers' distress and enhancement of the quality of care they provide.

Overall, informal caregivers evaluate the eLearning programs positively since they consider them useful and user-friendly, which is then reflected in their positive attitudes and empathy [16, 17]. This is mainly caused by the flexibility of these eLearning programs, their easy accessibility from anywhere and when needed (cf. [24]). Access to learning is one of the key factors influencing the growth of distributed learning environments. Users can access materials/resources at anytime and anywhere. Furthermore, they can proceed at their own pace. Thus, such an approach to learning enables a greater access to personalized learning and autonomy. Consequently, it means higher motivation and stimulation for its users.

However, as Moniz-Cook et al. [20] report, without ongoing review of implementation of recommended action plans, eLearning interventions are not that much effective at reducing challenging behavior in dementia when compared to usual care. In addition, the Chinese study [18] reveals that caregivers need an appropriate training since usually only younger caregivers are willing to use the eLearning programs, as well as those with higher educational level. Nevertheless, overall, there is a growing demand for the eLearning programs and resources, which would ensure quality and efficient services to residents and their families and provide nurses and caregivers with the necessary knowledge and skills to improve the quality of care (cf. [18]).

The limitations of this study consist in the lack of the research studies on this topic, follow-up observation and assessment period of the included publications. This might result in the overestimated effects of the use of eLearning as a valuable caregivers' support for patients in dementia, which can have a negative impact on the validity of these research studies.

5 Conclusion

The findings of this study indicate that the exploitation of eLearning as a support tool for caregivers in the management of dementia seems to be a promising method although its implementation requires the need of professional training of informal caregivers in the use of this technology. In addition, as Walsh [26] points out, eLearning technologies should be used purposefully and wisely in order to help deliver better care for patients and populations. For further research into the use of eLearning technologies and its consequences for human development and communication consequences see Pikhart [27–33].

Acknowledgments. This article is supported by the SPEV project 2019, run at the Faculty of Informatics and Management, University of Hradec Kralove, Czech Republic. The author thanks Josef Toman for his help with the data collection.

References

1. Dementia statistics. https://www.alz.co.uk/research/statistics
2. Corbett, A., Husebo, B.S., Achterberg, W.P., Aarsland, D., Erdal, A., Flo, E.: The importance of pain management in older people with dementia. Brit. Med. Bull. **111**(1), 139–148 (2014)
3. Craik, F., Salthouse, T.: The Handbook of Aging and Cognition, 2nd edn. Lawrence Erlbaum, Mahwah (2002)
4. Salthouse, T.: Consequences of age-related cognitive declines. Ann. Rev. Psychol. **63**, 201–226 (2012)
5. Hahn, E.A., Andel, R.: Non-pharmacological therapies for behavioral and cognitive symptoms of mild cognitive impairment. J. Aging Health **23**(8), 1223–1245 (2011)
6. De Vliegher, K., Aertgeerts, B., Declercq, A., Milisen, K., Sermeus, W., Moons, P.: Shifting care from hospital to home: a qualitative study. Pri. Health Care **25**(9), 24–31 (2015)
7. Alzheimer's Association: Alzheimer's Association Report 2015. Alzheimer's disease facts and figures. Alzheimer's & Dementia, vol. 11, pp. 332–384 (2015)
8. Wasilewski, M.B., Stinson, J.N., Cameron, J.I.: Web-based health interventions for family caregivers of elderly individuals: a scoping review (2017). http://dx.doi.org/10.1016/j.ijmedinf.2017.04.009
9. Klimova, B., Valis, M., Kuca, K.: Potential of mobile technologies and applications in the detection of mild cognitive impairment among older generation groups. Soc. Work Health Care **56**(7), 588–599 (2017)
10. Chong, M.C., Francis, K., Cooper, S., Abdullah, K.L., Thin, N., Hmwe, T., Sohod, S.: Access to, interest in and attitude toward e-learning for continuous education among Malaysian nurses. Nurse Educ. Today **36**, 370–374 (2016)
11. Delf, P.: Designing effective eLearning for healthcare professionals. Radiography **19**(4), 315–320 (2013)
12. Rider, B.B., Lier, S.C., Johnson, T.K., Hu, D.J.: Interactive web-based learning: translating health policy into improved diabetes care. Am. J. Prev. Med. **50**(1), 122–128 (2016)
13. Menard, P., Varpio, L.: Selecting an interprofessional education model for a tertiary health care setting. J. Interprof. Care **28**(4), 311–316 (2015)
14. Casimiro, L., MacDonald, C.J., Thompson, T.L., Stodel, E.: Grounding theories of W(e) Learn: a framework for online interprofessional education. J. Interprof. Care **23**(4), 390–400 (2009)
15. Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G.: The PRISMA group. Preferred reporting items for systematic review and meta-analysis: the PRISMA statement. PLoS Med. **6**(6), e1000097 (2009)
16. Hattink, B., Meiland, F., van der Roest, H., Kevern, P., Abiuso, F., Bengtsson, J., Droes, R. M.: Web-based STAR e-learning course increases empathy and understanding in dementia caregivers: results from a randomized controlled trial in the Netherlands and the United Kingdom. J. Med. Internet Res. **17**(10), e241 (2015)
17. Nunez-Naveira, L., Alonso-Bua, B., de Labra, C., Gregersen, R., Maibom, K., Mojs, E., Millán-Calenti, J.C.: UnderstAID, an ICT platform to help informal caregivers of people with dementia: a pilot randomized controlled study. BioMed. Res. Int. **2016**(5726465), 13 (2016)

18. Ho, D.W.H., Maka, V., Kwokab, T.C.Y., Au, A., Ho, F.K.Y.: Development of a web-based training program for dementia caregivers in Hong Kong. Clin. Gerontologist **38**, 211–223 (2015)
19. Wijma, E.M., Veerbeek, M.A., Prins, M., Pot, A.M., Willemse, B.M.: A virtual reality intervention to improve the understanding and empathy for people with dementia in informal caregivers: results of a pilot study. Aging Ment. Health (2017). https://doi.org/10.1080/13607863.2017.1348470
20. Moniz-Cook, E., Hart, C., Woods, B., Whitaker, C., James, I., Russell, I., Manthorpe, J.: Challenge Demcare: management of challenging behaviour in dementia at home and in care homes – development, evaluation and implementation of an online individualised intervention for care homes; and a cohort study of specialist community mental health care for families. Programme Grants Appl. Res. **5**(15) (2017). https://doi.org/10.3310/pgfar05150
21. Boots, L.M.M., Vugt, M.E., Knippenberg, R.J.M., Kempen, G.I.J.M., Verhey, F.R.J.: A systematic review of internet-based supportive interventions for caregivers of patients with dementia. Int. J. Geriatr. Psychiatry **29**(4), 331–344 (2014)
22. Dam, A.E., de Vugt, M.E., Klinkenberg, I.P., Verhey, F.R., van Boxtel, M.P.: A systematic review of social support interventions for caregivers of people with dementia: are they doing what they promise? Maturitas **85**, 117–130 (2016)
23. Kurz, A., Bakker, C., Böhm, M., Diehl-Schmid, J., Dubois, B., Ferreira, C., de Vugt, M.: RHAPSODY – internet-based support for caregivers of people with young onset dementia: program design and methods of a pilot study. Int. Psychogeriatr. **28**(12), 2091–2099 (2016)
24. Tannebek, K., Hultberg, P., Refsgaard, E., Hansen, B.H., Waldemar, G.: The Danish ABC dementia: new e-learning for professional caregivers. Alzheimer's Dement. **10**(4), P180–P181 (2014)
25. Moreno, P.A., Garcia-Pacheco, J.L., Charvill, J., Lofti, A., Langensiepen, C., Saunders, A., Gomez, E.J.: iCarer: AAL for the informal carers of the elderly. Stud. Health Technol. Inf. **210**, 678–680 (2015)
26. Walsh, K.: The future of e-learning in healthcare professional education: some possible directions. Ann. Ist. Super. Sanita. **50**(4), 309–310 (2014)
27. Pikhart, M.: Current intercultural management strategies. The role of communication in company efficiency development. In: Proceedings of the 8th European Conference on Management, Leadership and Governance, pp. 327–331 (2012)
28. Pikhart, M.: Communication based models of information transfer in modern management – the use of mobile technologies in company communication. Innov. Manage. Educ. Excellence Vis. **2020**, 447–450 (2018)
29. Pikhart, M.: Electronic managerial communication: new trends of intercultural business communication. In: Innovation Management and Education Excellence Through Vision 2020, IBIMA 2018, pp. 714–717 (2018)
30. Pikhart, M.: Managerial communication and its changes in the global intercultural business world. In: Web of Conferences, vol. 26, pp. 1–5 (2016)
31. Pikhart, M.: Intercultural linguistics as a new academic approach to communication. In: Web of Conferences, vol. 26, pp. 6–9 (2016)
32. Pikhart, M.: Implementing new global business trends to intercultural business communication. Procedia Soc. Behav. Sci. **152**, 950–953 (2014)
33. Pikhart, M.: New horizons of intercultural communication: applied linguistics approach. Procedia Soc. Behav. Sci. **152**, 954–957 (2014)

Mobile Learning and Its Impact on Learning English Vocabulary

Blanka Klimova(✉)

Faculty of Informatics and Management, University of Hradec Kralove,
Rokitanskeho 62, 50003 Hradec Kralove, Czech Republic
blanka.klimova@uhk.cz

Abstract. Mobile learning is nowadays becoming a well-established methodology in the institutions of higher learning. The key reasons comprise easy portability and accessibility of mobile devices, as well as ubiquity of wireless environment, and the opportunity of informal and personalized approaches to learning. The purpose of this article is to discuss whether the use of a mobile application developed for English vocabulary learning for students of Management of Tourism at the Faculty of Informatics and Management in Hradec Kralove, Czech Republic, is effective. The methodology of this study was based on an experiment, which consisted of a group of students who used the app and of those who did not. The findings of this study show that the use of mobile app for English vocabulary learning is effective as suitable support to traditional, face-to-face classes.

Keywords: Mobile app · Students · English · Use · Effectiveness

1 Introduction

Currently, mobile technology is an inseparable part of almost all human activities, including education. In fact, mobile learning is nowadays becoming a well-established methodology in the institutions of higher learning. The key reasons comprise easy portability and accessibility of mobile devices, as well as ubiquity of wireless environment, and the opportunity of informal and personalized approaches to learning [1–4]. Mobile learning is mainly operated via mobile phones, respectively smartphones as statistics shows [5].

No wonder that mobile devices are also widely exploited in English language learning. Such language learning with the help of mobile devices, especially smartphones, is called Mobile Assisted Language Learning (MALL) [6]. It is considered to be part of the so-called Computer Assisted Language Learning (CALL) [7]. However, as Kukulska-Hulme and Shield [8] emphasize, MALL is different from CALL because it emphasizes spontaneity of access and interaction across different contexts of use.

For learning any language, vocabulary learning is essential, specifically for language proficiency. The development of all language skills, i.e., listening, reading, writing, and speaking, depends on the range of vocabulary one possesses (cf. [9]). Therefore, any method, which can improve student's retention of new words and phrases is sought by language practitioners. One of these methods is also MALL. The

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 271–276, 2020.
https://doi.org/10.1007/978-981-32-9244-4_38

effectiveness of mobile devices in English vocabulary learning has been confirmed by a number of recent research studies [10–16]. For example, Mahdi [10] in his meta-analysis proved that mobile devices could be effective in learning new vocabulary. The findings of his research study revealed that the use of mobile devices in vocabulary learning had been more effective than the traditional way of teaching. Interestingly, the findings showed that adult learners could benefit from the use of mobile devices in learning vocabulary more than their younger counterparts. In addition, Elaish et al. [17] claim that pure mobile applications are the most widely-used applications in the English m-learning context.

The purpose of this article is to discuss whether the use of a mobile application developed for English vocabulary learning for students of Management of Tourism at the Faculty of Informatics and Management in Hradec Kralove, Czech Republic, is effective. The results represent a comparison of two semester's English vocabulary and phrases learning via a mobile app.

2 Materials and Methods

2.1 Participants

There were two sets of students. They consisted of students of Management of Tourism in their third year of study using a mobile app in the winter semester 2017 and 2018. All were full-time students and used the mobile app as support to traditional, face-to-face classes.

In the winter of 2017, 31 students were enrolled in the course on practical English. Out of these students, 19 students used the mobile app and 12 students did not. In the winter semester of 2018, altogether 25 students were enrolled in the course; 19 students used the app and six did not. All the contact classes lasted 90 minutes per week and were held from the end of September to the mid of December.

2.2 Study Design

The methodology of this study was based on an experiment, which consisted of a group of students who used the app and of those who did not. The content of the mobile app was tailored to students' needs. The needs analysis was carried out at the first lesson of the winter semester. Thus, some of the new words and phrases were modified for the winter semester of 2018 because students' needs differ from one individual to another and from one year to another.

The mobile app was developed already before the winter semester of 2017. It is called *Angličtina* (English) *TODAY*. It is a software architecture that uses the server part, the web application and the mobile application [3]. The mobile app was developed in the original design only for the Android operating system. The reason was a larger percentage of students using this operating system. In fact, most users worldwide use this operating system (cf. [18]). The application proved to be beneficial after annual use [4]. In order to offer this solution to other students, a mobile application available for the iOS operating system was developed one year later. This was performed before the

winter semester of 2018. Thus, nowadays, students can exploit both operating system when using the mobile app. Figure 1 illustrates the teacher's interface and Fig. 2 a mobile app screen.

Fig. 1. Teacher's interface [20].

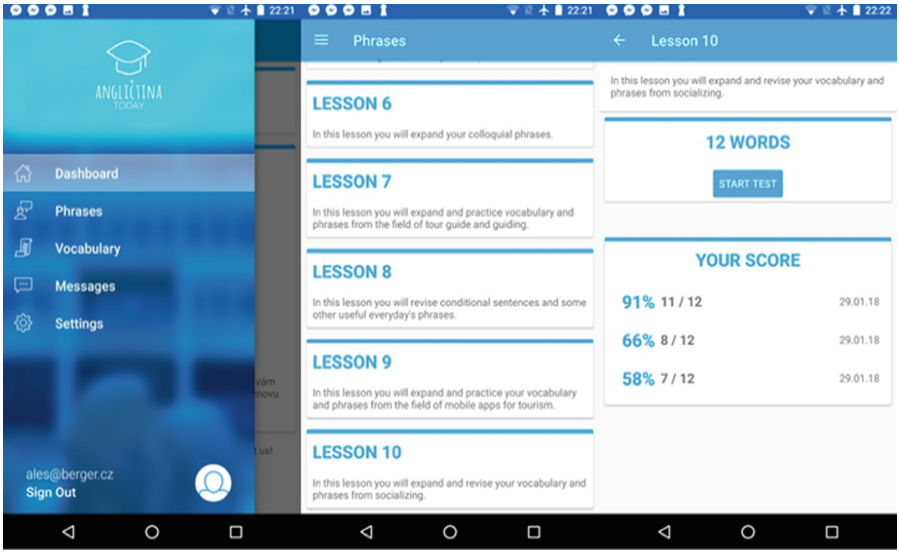


Fig. 2. Mobile app screen.

Altogether, there are ten parallel lessons of vocabulary and phrases. The content of each lesson is physically completed with words and phrases by the teacher. The students must translate the word or the phrase from their native language into English. Each lesson is done as a test and comprises on average of 15-18 new words and 10 new phrases. The selected words and phrases are always those discussed in the face-to-face classes so that the students who do not use the mobile app can write them down and practice then from their notes. All the words and phrases are related to their field of study, i.e., tourism, the topics generated by the needs analysis at the beginning of the course. Furthermore, the teacher tries to encourage the students to revise and practise the new words and phrases via notifications sent to them on their smartphones at least twice a week.

3 Results and Discussion

In order to prove the effectiveness of the use of mobile app on students' vocabulary retention, the results of students' achievement tests and their attendance in the regular, face-to-face classes were used. The pass mark for doing the final achievement test was 50%, i.e., 30 points. Students had to attend 70% of classes at minimum. Table 1 below illustrates the results.

Table 1. An overview of students' results with and without the mobile app and attendance.

	No. of students with mobile app	No. of students without mobile app	No. of students attending 70% of the course	No. of students not attending 70% of the course	No. of students who passed the final test	No. of students who did not pass the final test	No. of students who passed the final test and used the app	No. of students who did not pass the final test and did not use the app
2017	19	12	13	18	24	7	23	6
2018	19	6	12	13	18	7	16	4

Thus, the results in Table 1 show that the majority of those who were using the mobile app succeeded in passing the test. Furthermore, the results also indicate that those who were not attending the face-to-face classes, but were using the mobile app passed the test as well. In the year of 2017, there were four students who neither attended the regular classes not used the app, but they eventually passed the test. The author of this article might just assume that these students might have consulted the materials with those who attended the class. In the winter semester of 2018, this issue did not repeat.

Therefore, the findings of this study confirm the results of other studies on the effective use of mobile app in English vocabulary learning, such as [10] or [16].

One of the benefits of the designed mobile app *Angličtina TODAY* [19] is also its personalized approach and the possibility of continuous modification. Another benefit is also revision of the most problematic words, which were identified by mobile app scores. All this might then increase students' motivation to use the app and practise the new words and phrases (cf. [20]).

For further utilization of mobile apps and the consequences it can bring in learning process and communication in the global world see research of Pikhart [21–25].

4 Conclusion

The findings of this study show that the use of mobile app for English vocabulary learning is effective as suitable support to traditional, face-to-face classes. Future research should continue in order to obtain more subject samples and thus, more conclusive results. In addition, the author might research students' motivation and attitude towards the use of mobile app.

Acknowledgments. This article is supported by the SPEV project 2019, run at the Faculty of Informatics and Management, University of Hradec Kralove, Czech Republic. The author thanks Ales Berger for his help with the data collection.

References

1. Klimova, B.: Mobile phones and/or smartphones and their apps for teaching English as a foreign language. Educ. Inf. Technol. **23**(3), 1091–1099 (2017)
2. Mehdipour, Y., Zerehkafi, H.: Mobile learning for education: benefits and challenges. Int. J. Comput. Eng. Res. **3**(6), 93–101 (2013)
3. Berger A., Klímová B.: Mobile application for the teaching of English. In: Park, J., Loia, V., Choo, K.K., Yi, G. (eds.) Advanced Multimedia and Ubiquitous Engineering, MUE 2018, FutureTech 2018. LNEE, vol. 518, pp. 1–6 (2019)
4. Klímová, B., Berger, A.: 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, vol. 11284, pp. 3–11 (2018)
5. Mobile learning. https://www.shiftelearning.com/blog/bid/331987/mobile-learning-stats-that-will-make-you-rethink-your-training-strategy
6. Chinnery, G.: Going to the MALL: mobile assisted language learning. Lang. Learn. Technol. **10**(1), 9–16 (2006)
7. Yang, J.: Mobile assisted language learning: review of the recent applications of emerging mobile technologies. Engl. Lang. Teach. **6**(7), 19–25 (2013)
8. Kukulska-Hulme, A., Shield, L.: An overview of mobile assisted language learning: from content delivery to supported collaboration and interaction. ReCALL **20**(3), 271–289 (2008)
9. Lewis, B.: The Lexical Approach. LTP, London (1993)
10. Mahdi, H.S.: Effectiveness of mobile devices on vocabulary learning: a meta-analysis. J. Educ. Comput. Res. **56**(1), 134–154 (2017)
11. Wu, Q.: Learning ESL vocabulary with smartphones. Procedia – Soc. Behav. Sci. **143**, 302–307 (2014)

12. Wu, Q.: Designing a smartphone app to teach English (L2) vocabulary. Comput. Educ. **85**, 170–179 (2015). https://doi.org/10.1016/j.compedu.2015.02.013
13. Song, M., Chen, L.: A review on English vocabulary acquisition and teaching research in recent 30 years in China. Sci. J. Educ. **5**(4), 174–180 (2017)
14. Niangah, T.M., Nezarat, A.: Mobile-assisted language learning. IJDPS **3**(1), 309–319 (2012)
15. Wang, B.T.: Designing mobile apps for English vocabulary learning. Int. J. Inf. Educ. Technol. **7**(4), 279–283 (2017)
16. Basal, A., Yilmaz, S., Tanriverdi, A., Sari, L.: Effectiveness of mobile applications in vocabulary teaching. Contemp. Educ. Technol. **7**(1), 47–59 (2016)
17. Elaish, M.M., Shuib, L., Ghani, N.A., Yadegaridehkordi, E., Alaa, M.: Mobile learning for English language acquisition: taxonomy, challenges, and recommendations. IEEE Access **5**, 19033–19047 (2017)
18. Statcounter. http://gs.statcounter.com/os-market-share/mobile/worldwide/#monthly-200901-201812
19. Angličtina Today. https://www.anglictina.today/
20. Goz, F., Ozcan, M.: An entertaining mobile vocabulary learning application. In: EPESS, vol. 7, pp. 63–66 (2017)
21. Pikhart, M.: Current intercultural management strategies. The role of communication in company efficiency development. In: Proceedings of the 8th European Conference on Management, Leadership and Governance (ECMLG), pp. 327–331 (2012)
22. Pikhart, M.: Communication based models of information transfer in modern management – the use of mobile technologies in company communication. In: Innovation Management and Education Excellence through Vision 2020, IBIMA 2018, pp. 447–450 (2018)
23. Pikhart, M.: Electronic managerial communication: new trends of intercultural business communication. In: Innovation Management and Education Excellence through Vision 2020, IBIMA 2018, pp. 714–717 (2018)
24. Pikhart, M.: Implementing new global business trends to intercultural business communication. Procedia Soc. Behav. Sci. **152**, 950–953 (2014)
25. Pikhart, M.: New horizons of intercultural communication: applied linguistics approach. Procedia Soc. Behav. Sci. **152**, 954–957 (2014)

Improvement of the User Throughput with Relaying Through CR-Based UAV

Suho Choi, Wooyeob Lee, and Inwhee Joe(✉)

Department of Computer and Software, Hanyang University, Seoul, Korea
{aiek26l2, matiasl2, iwjoe}@hanyang.ac.kr

Abstract. In a wireless communication environment, using UAV to relay network can improve user throughput. This is because the UAV's mobility can maintain the connection of the user's network and extend the communication range. This is useful when obstacles such as mountains or large buildings degrade the quality between the base station and the user or the distance between the base station and the user is long. Moreover, the cognitive radio technology allows more users to use the band of licensed users who are not using it to relay the network. In this paper, UAV relays using licensed user's bandwidth through Cognitive Radio (CR) technology and simulated to improve the throughput of users located far from the base station. Simulation results show that user throughput is improved.

Keywords: Cognitive radio · Communication relay · UAV

1 Introduction

In a wireless communication environment, the relay of the network through the UAV is attractive. The reason for this is that the mobility of the UAV maintains the user's network continuity in obstacles such as mountains or large buildings. Furthermore, the communication range of the user can be expanded [1, 2]. However, network relay cannot be used by many users because of its narrow bandwidth.

On the other hand, wireless communication uses limited frequency resources and therefore requires national management. Generally, each user is licensed and uses a fixed frequency allocation policy. However, with the spread of various devices using wireless communication, the frequency resources available for existing policies are being exhausted.

Mitola has proposed a CR technology that allows opportunistic use of the frequency band of unlicensed licensees [3]. CR technology detects non-licensed users' usage of the licensee's frequency. If the license user's bandwidth is clear, the non-licensed user transfers the data over that frequency band. Therefore, a limited frequency can be efficiently used.

This paper is organized as follows. In Sect. 2, introduces the related works, and Sect. 3 explains our proposed model. We present procedures for UAV to initiate network relay. The simulation setup and results are presented in Sect. 4, followed by conclusions in Sect. 5.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 277–281, 2020.
https://doi.org/10.1007/978-981-32-9244-4_39

2 Related Works

Recently, communication relay using UAV has been actively studied. Technological advances enable miniaturization and performance enhancement of sensors and batteries can provide significant benefits in mobile ad hoc networks (MANETs) [4]. Moreover, UAV can extend the typical communication limit distance. This feature is effective for communication between the vessels using only limited satellite communications [5].

On the other hand, there is a problem in using a cellular network system in a UAV. The manufacturer must be in contact with the carrier and the user must pay the communication fee. Moreover, when Wi-Fi is used, there is a communication range restriction with a limited output. And interference problems with devices using the same channel occur [6].

3 Proposed Model

We used CR technology to relay, including the frequency band of the licensee (Primary User, PU). Figure 1 is a block diagram of UAV using CR. We receive the signal using an RF receiver block with energy detection. The analysis block then analyzes and lists the available frequency bands for the UAV. In this way, we have taken the threshold criterion that the PU does not occupy the frequency band. In the Control block, it communicates with the BS based on the list passed in the Analysis block.

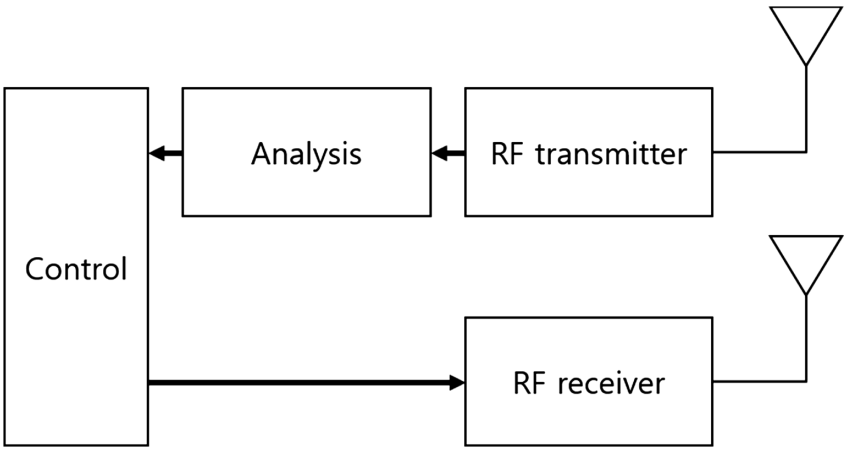


Fig. 1. CR-UAV block diagram

The following example shows the mechanism process (Fig. 2). The UAV specifies the frequency band of the available PU through the list obtained through the analysis block. The data is transmitted and received through the corresponding frequency band. Periodic sensing can confirm that the PU has changed from Idle to wake-up state. At the same time, check some of the lists to be able to jump to another PU. If there is no

frequency band that can be occupied, it is sensed again after having a certain back-off time.

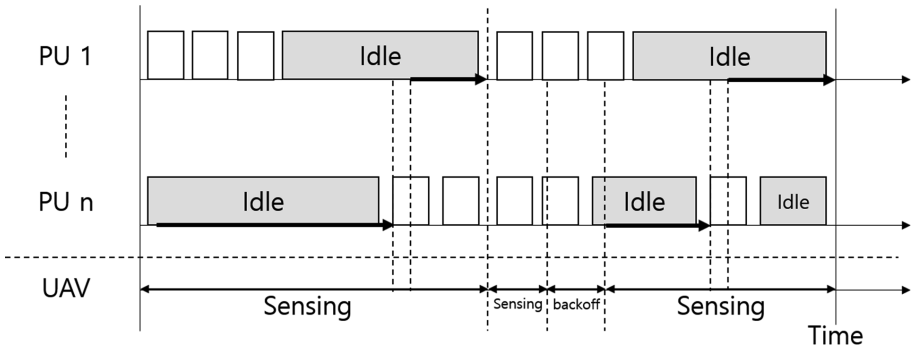


Fig. 2. Examples of CR mechanisms

4 Simulation

For the simulation, we used OPNET version 14.5 A. Our network configuration is set and compared as shown in Fig. 3. The simulation background is the campus and set at 10 × 10 (km). The parameters considered for the simulation are shown in Fig. 4. The bandwidth of the UAV has been assigned 20 by default. However, I set it to be able to expand without specifying the maximum value. The simulation was run for 30 s.

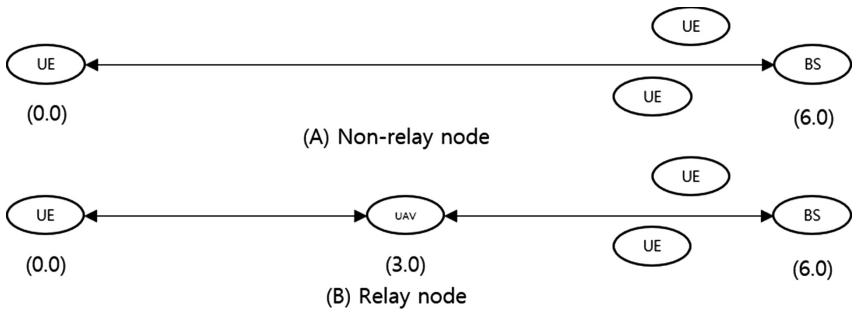


Fig. 3. Network configuration for simulation

The simulation results are shown in Fig. 5. The graph compares the average receiver throughput of the UE located at (0.0) in Fig. 1. In the graph, the throughput of the *relay-node* environment is more than twice that of the *non-relay-node* environment. this is because in the case of a *non-relay-node* environment, the distance between the UE and the BS is too long, so that many packets transmitted from the BS have been lost to the UE.

	UE	UAV	BS
Tx Power(dBm)	20		35
Bandwidth(kHz)	20		60
Frequency(MHz)	2,600		
Data Rate(bps)	1,024		
Transmission mode	2 x 2 MIMO		
Path loss Model	Free-space path loss model		

Fig. 4. Node parameter for simulation

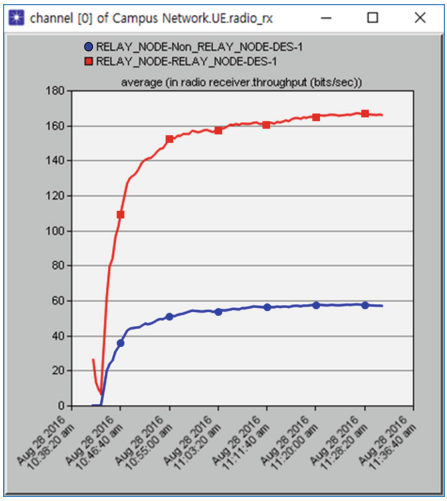


Fig. 5. Average Rx throughput of UE

On the other hand, in the *relay-node* environment, the UAV relays the packet transmitted from the BS to the UE. Moreover, the UAV senses the frequency band of neighboring nodes and extends the bandwidth through the clear band. Therefore, the UE performance is greatly improved by transmitting more packets.

Figure 6 shows the average TX throughput graph of BS in each environment model. In the graph, the throughput of the two models is not much different. This is because the BS transmits packets at the same bandwidth. We have verified the operation of the proposed model through these results.

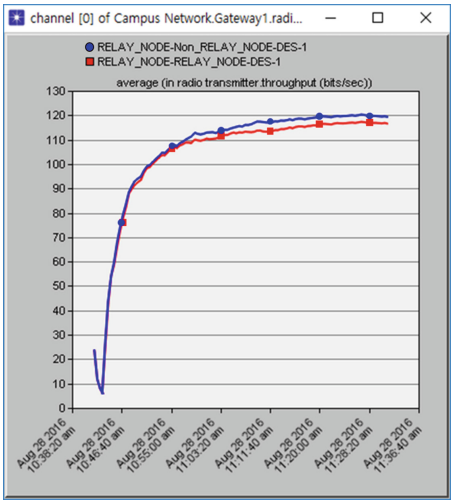


Fig. 6. Average Tx throughput of BS

5 Conclusion

In this paper, we have simulated the effect of communication relay over CR based UAV. Simulation results show that the network relay improves the throughput of users located far from the base station. Further, it was confirmed that the result was obtained through surplus resources without adding the resources of the base station.

Acknowledgement. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science, ICT & Future Planning) (NO. 2016R1A2B4013118).

References

1. Cheng, C., Hsiao, P., Kung, H., Vlah, D.: Maximizing throughput of UAV-relaying networks with the load-carry-and-deliver paradigm. In: IEEE WCNC, pp. 4417–4424 (2007)
2. Basu, P., Redi, J., Shurbanov, V.: Coordinated flocking of UAVs for improved connectivity of mobile ground nodes. In: IEEE MILCOM, vol. 3, pp. 1628–1634 (2004)
3. Mitola, J., Maguire, G.Q.: Cognitive radio: making software radios more personal. IEEE Pers. Commun. **6**, 13–18 (1999)
4. Han, Z., Swindlehurst, A.L., Liu, K.J.R.: Optimization of MANET connectivity via smart deployment/movement of unmanned air vehicles. IEEE Trans. Veh. Technol. **58**, 3533–3546 (2009)
5. Seung, K., Jin, S.: Efficient communication relay strategy using small UAVs. In: The Korea Society for Aeronautical and Space Sciences, pp. 389–393 (2012)
6. Seung, S., Jin, K., Kyung, P.: Overview and issues of drone wireless communication. J. Korea Inst. Commun. Sci. **33**(2), 93–99 (2016)

Distribution Pattern Learning for Social Resources in a Golden Snub-Nosed Monkey Community

Qiang Hu, Xiaoqian Mi, Songtao Guo, Huawei Liu, Fengyi Song, Kun Yang, and Xiaoyan Yin(✉)

School of Information Science and Technology, Northwest University, Xi'an 710127, China
{huqiang, mxq, songfengyi, ykun}@stumail.nwu.edu.cn, songtaoguo@nwu.edu.cn, l5609282903@163.com, SCxiaoyanyin@gmail.com

Abstract. The law of survival of the fittest and struggle for existence of wildlife in nature is closely related to food. We can imagine that lock of food is the root cause of wildlife endangerment. Thus, it would be very meaningful if the monitoring data could confirm the importance of food for survival. To illustrate functioning rules of the golden snub-nosed monkey society and the importance of food, we explore the behavior data of 12 golden snub-nosed monkey families and then adopt tensor decomposition to dig out the potential relationship between the social class and food allocation. Fortunately, the results of tensor decomposition are consistent with the observations of animal protection experts.

Keywords: Resource allocation · Golden snub-nosed monkey · Distribution pattern · Tensor decomposition

1 Introduction

In the long-term survival and evolution process, animals often face the problem of food shortage, especially in bad weather days. To face the challenge of food, there are both competition and cooperation among wild animals, and a fixed organizational structure is gradually formed in a golden snub-nosed monkey community. Clustering is a phenomenon that occurs in different modes for most Primates [1]. However, survival of the fittest and natural selection of things lead to the division of social hierarchy and the continuous reorganization of clusters, and stronger individuals get more and better resources, including food and spouses. By studying the potential relationship between resource allocation and the social class in the golden snub-nosed monkey society, the resource allocation pattern of the golden snub-nosed monkey society is figured out, and the formation and evolution of the primate social system are revealed.

Ranking of individuals is based on one certain criterion, and the social class can reflect the differences of individuals in competing for resources. When resources are scarce, higher social class can get enough resources. When resources are abundant, the

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 282–291, 2020.
https://doi.org/10.1007/978-981-32-9244-4_40

higher social class can get the priority of distribution. In the wildlife world, differentiation of social classes is mostly based on the fighting power of animals, and stronger individuals always correspond to higher social class.

To reveal the formation and evolution rules of Primate, understand the changes of competition and cooperation mode and community structure of basic units of the system under different food supply, we explore the golden snub-nosed monkey with special structure in Qinling Mountains. The basic unit (family interchangeably) of golden snub-nosed monkey society is the One-Male Unit (OMU), which consists of an adult male and several adult females, as well as some sub-adult and adolescent monkeys and baby monkeys, and the social class of the unit is determined by the fighting power of the male monkey. The golden snub-nosed monkey is active in Qinling, Shaanxi, Baihe, Sichuan, Shennongjia, Hubei, Lushan and Gansu. However, the living environment and habitat of Sichuan golden snub-nosed monkey has been widely destroyed, and has been listed as the world's endangered species. Experts believe that studying the resource allocation pattern of the golden snub-nosed monkey society is the most effective way to protect them.

In order to help experts explore the golden snub-nosed monkey society, the monitoring data of 12 golden snub-nosed monkey families is used in this paper, which comes from an artificial feeding site of the golden snub-nosed monkey in Qinling Mountains. The feeding site provides enough food to golden snub-nosed monkeys twice a day, one in the morning and the other in the noon. Whenever a fixed feeding time comes, the monkeys jump to the feeding site in crowds and groups. After naming each monkey, the recorder writes down the changes of the members of each family, the results of each fighting, the number of winnings or defeats, the order of eating and the area occupied by families during eating.

Based on tensor decomposition, this paper studies the potential relationship between the social class and resource allocation. By using the behavior data of 12 golden snub-nosed families, we study the potential relationship among the size of units, eating order and the areas occupied by families during eating, as well as analyze the potential relationship among the fighting power (i.e., the number of winnings/defeats) of the family, the size of units and eating order. The results of the study are consistent with those observed by animal protection experts.

In this paper, we have made the following contributions:

(1) We adopt tensor decomposition to dig out the potential relationship between resource allocation and the social class for the golden snub-nosed monkey society.
(2) We use HOSVD (high-order singular value decomposition) for matrix decomposition to solve the problem of data sparseness and information missing in data sets.
(3) We verify that the results of tensor decomposition are consistent to observations of golden snub-nosed monkey protection experts.

2 Related Work

Analysis of Wildlife Behavior. In order to protect wildlife effectively, it is necessary to conduct research on animal life and behavior. In [2], a few observations which are about the role of dominance of infra-human Primates were shared. Maslow [3] extended the concept of hierarchical structure to more complex primate species. Wittemyer et al. [4] studied African elephants, and analyzed the hierarchical dominance structure and social organization. In [5], He et al. assessed the social relationships among adult females and clarified possible explanations for the strategies of female mating competition in the one-male units. The research on the social hierarchical relationship and resource allocation system of Qinling golden monkeys mainly focuses on the predation behavior and attacking behavior of Qinling golden monkeys [6]. As time went by, scientists use more behavioral indicators to study the hierarchical relationship of non-human primates, such as mating behavior [7]. Feeding behavior mainly studies the order and scope of feeding [8]. The possession of resources is usually another indicator for determining the social hierarchy and resource allocation of the Qinling Golden Monkey.

Tensor Decomposition. Tensor decomposition has been studied extensively and applied to various fields, signal processing, numerical analysis, data mining, linear algebra [9]. In [10], Ailon et al. proposed a novel dimension reduction scheme for web queries, while avoiding the curse of dimensionality. Yao et al. [11] proposed a collaborative filtering method based on nonnegative tensor factorization that exploits a high-order tensor, instead of traditional user-location matrix to model multi-dimensional contextual information. In [12], tensor decomposition is leveraged to achieve position-sensitive over-the-counter drug recommendations. In [13], Tang et al. proposed a tri-clustered tensor completion framework to refine social image tags. To improve users' quality of experience, Hu et al. [14] prefetched videos based on tensor factorization. However, none of these frameworks can be directly applied in this paper.

3 Our Proposed Algorithm

In this section, we first introduce the tensor decomposition model, which can decompose a sparse tensor into a core-tensor and the factor matrices to mine the potential properties of three dimensions. Then, we apply HOSVD Tensor Decomposition to reveal the latent relationships among the size of units, eating order and the areas occupied by families during eating, as well as analyze the potential relationship among the fighting power (i.e., the number of winnings/defeats), the size of units and eating order.

3.1 Tensor Decomposition Model

The tensor is a multidimensional matrix that can store multidimensional data. A first-order tensor is a vector, a second-order tensor is a matrix, and tensors of order three or higher are called higher-order tensors, as shown in Fig. 1. The third-order tensor has three indices, can describe the relationship between data in three dimensions.

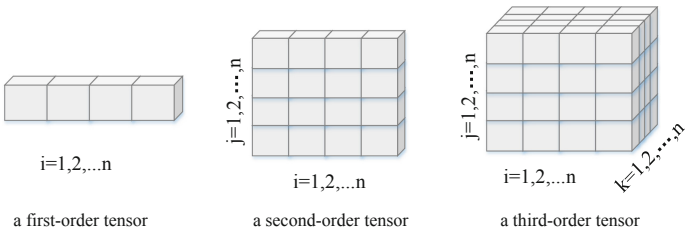


Fig. 1. Illustrations of three kinds of tensors

The high-order singular value decomposition method can reduce the dimensionality of the three-dimensional tensor, and explore the potential relationship between the three-dimensional tensors. The basic tensor decomposition model can be regarded as a high-order extension of matrix singular value decomposition, which is Candecomp Parafac decomposition and Tucker decomposition. The Candecomp Parafac (CP) decomposition [15] of the third-order tensor can be defined as follow.

$$x \approx \sum_{r=1}^{R} a_r \circ b_r \circ c_r \tag{1}$$

where $\circ$ represents the outer product of the vector, R is the rank of the tensor, and a_r, b_r, c_r is the unit vector.

Tucker decomposition is a high-order form of principal component analysis. Traditional principal dimension analysis methods such as PCA and LDA directly reduce the dimension to a two-dimensional matrix when dealing with high-dimensional data, which has the integrity of missing information. Tucker decomposition is to represent the n-order tensor $x \in R^{I_1 \times I_2 \cdots I_N}$ as a modular multiplication of a core tensor and an n-dimensional factor matrix. Each factor matrix is called the principal component of the tensor in this dimension. This method preserves the original tensor information as much as possible. The Tucker decomposition of the third-order tensor is defined as follows.

$$x \approx S \times_A A \times_B B \times_C C = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} S_{pqr} a_p \circ b_q \circ c_r \tag{2}$$

where S is the core tensor, and A, B, and C are the factor matrices of three dimensions, respectively.

3.2 Distribution Pattern Learning

We formulate ranking of area occupied by 12 families under different eating orders as a three-order tensor $A \in R^{M \times T \times L}$, where **M** is the number of sizes of families, **T** is the order of eating, and **L** is the ranking of area occupied by families during eating. Each entry $a_{i,j,l} \in A$ represents the size of a family m to the number of ranking of the area l in the eating order t.

(1) HOSVD Tensor Decomposition: to reveal the latent semantic associations between the three components (golden monkey, feeding order, feeding range) of tensor A, we apply the HOSVD decomposition as shown in Fig. 2, which can decompose a sparse tensor into a core-tensor and the factor matrices to mine the potential properties of golden monkey, eating order, area occupied by families during eating.

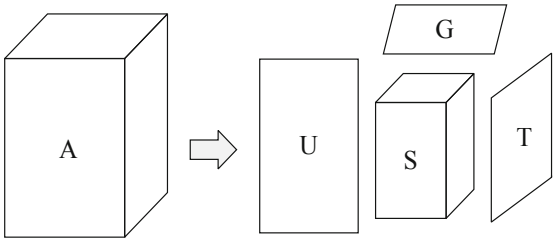


Fig. 2. HOSVD 3-order tensor decomposition.

(2) Mode-n Expansion: based the initial tensor, we expand it along mode-n to get three two-dimensional matrices **A1**, **A2** and **A3**. Mode-n expansion is shown in Fig. 3.

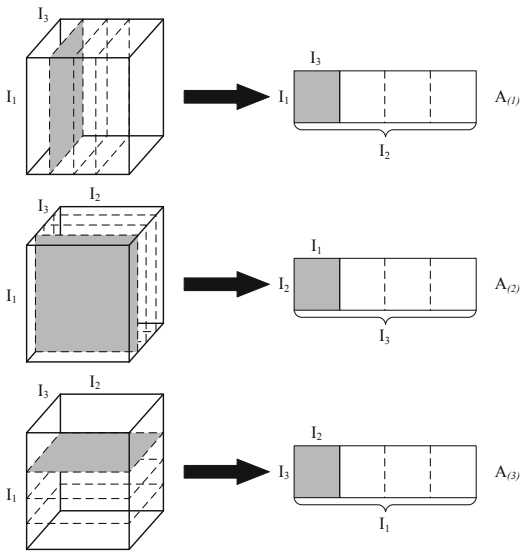


Fig. 3. Mode-n expansion

(3) Singular value decomposition: the singular value decomposition of A1, A2 and A3 is performed to obtain the transposed matrices of the left singular matrix U1, U2, U3, the right singular matrix V1, V2, V3, as well as the singular diagonal matrix W1, W2, W3.
(4) Approximation: select the ideal approximation and keep the initial diagonal matrix W1, W2, W3.
(5) The core tensor construction: to obtain the core tensor, we have

$$S = A \times_1 U_{C_1}^{(1)^T} \times_2 U_{C_2}^{(2)^T} \times_3 U_{C_3}^{(3)^T} \tag{3}$$

(6) The target tensor construction: based on the core tensor, the target tensor is calculated as follows.

$$\widehat{A} = S \times_U U \times_G G \times_T T \tag{4}$$

Based on the results of HOSVD, the approximate tensor can fill the sparse and unknown element in the initial tensor, and we can obtain the latent relationship between sizes of families, eating order, area occupied by families during eating.

3.3 The Latent Relationship Learning

Similar to Sect. 3.2, we apply tensor decomposition to study the relationship between sizes of families, eating ordering and fighting power ranking of families. We firstly construct the initial three-dimensional tensor $A \in R^{N \times P \times Q}$, where N is the number of sizes of families, P is the order of eating, and Q is the number of winnings or defeats. Secondly, based on the HOSVD decomposition, we can reconstruct the approximate tensor $\widehat{A}$. The approximate tensor can fill the unknown element in the initial tensor, and we can explore the potential relationship between sizes of families, eating ordering and fighting power ranking of families.

4 Performance Evaluation

In this section, we first give a detailed description of the dataset used in this paper. Then, we present the experimental results and experimental explanations.

4.1 Dataset

We obtained behavior data of 12 golden snub-nosed monkey families from October 14, 2015 to December 11, 2015, including the number of family members, eating order and fighting data. We can identify individual monkeys, and name 12 families as BX, SX, SQ, LD G3, WX, WF, ZB, ST, TB, BG and ZD, respectively. The eating data recorded the time, sequence of each family entering the feeding site, and area occupied by each family. Fighting data mainly record the initiator and start time of unfriendly behavior, receiver and receive time, and fighting results. Aggressive behavior includes biting,

beating, catching up and replacing, as well as yielding behavior includes avoidance and escape. In addition, individuals involved in joint attacks at the time of conflict also are recorded.

4.2 The Results of Distribution Pattern Learning

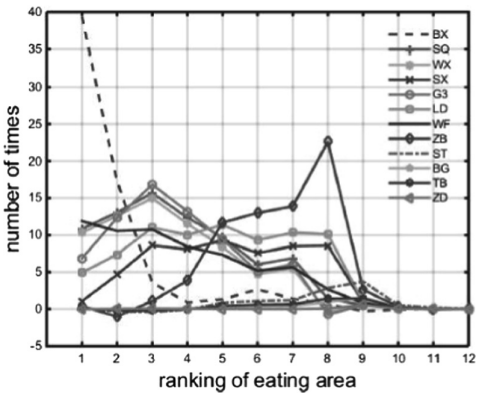
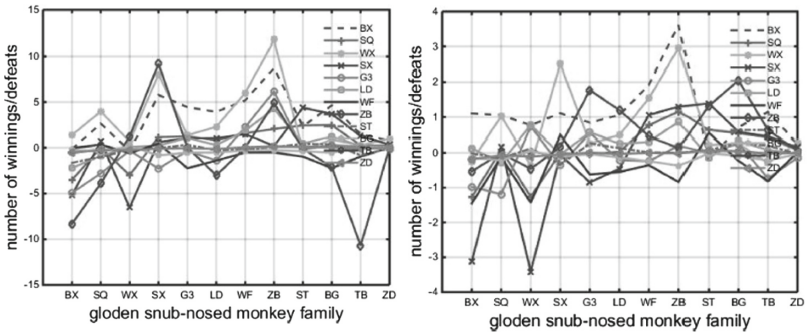


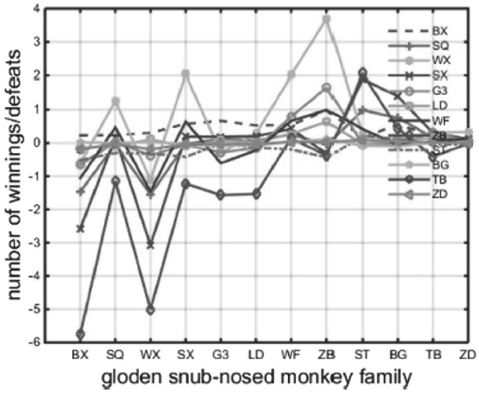
Fig. 4. Ranking comparison of the area occupied by the first sequence of families

Based on the behavior data of 12 golden snub-nosed monkey families, taking the first sequence of eating as an example, we analyzed the relationship between the number of golden monkey family members, eating order and the area occupied by families during eating, as shown in Fig. 4. Every family has the opportunity to enter the feeding site in the first sequence, but the number of times is different, and the area occupied is also different. The frequency of BX, SX, SQ, LD G3, WX and WF was significantly higher than that of ZB, ST, TB, BG and ZD. Relatively speaking, the higher the eating order, the larger the area. The eating order is positively correlated to the area. The number of monkey family members implies that the fighting power of the male monkey. The stronger the fighting power of the male monkey, the greater the attraction to female monkeys, the larger the family, the more food grabbed, the more family members, and the higher the social class. From Fig. 4, we can know that the frequency of family BX is significantly higher than that of other families, and the frequency of family ZD is lower than that of other families. Therefore, we can inferred that BX belongs to the highest social class and ZD belongs to the lowest social class.

Similarly, we analyzed the relationship between the eating order of families and fighting power of the male monkeys. As shown in Fig. 5, the lower the eating order, the lower the fighting power of the monkey family. BX, SQ and WX have higher number of winnings, and BX has the highest rank. The defeat times of ZB, ST, BG, TB and ZD are higher than that of other families. Therefore, we can inferred that BX belongs to the highest social class, while ZB, ST, BG, TB and ZD belong to the lowest social class.



(a) The sequence of eating is 1 (b) The sequence of eating is 2



(c) The sequence of eating is 3

Fig. 5. Illustrations of the relationship between eating order and fighting power for 12 golden snub-nosed monkey families.

4.3 Performance Analysis

Based on the results of tensor decomposition, we find that the social class of BX, SX, SQ, LD, G3, WX and WF should be higher than that of ZB, ST, TB, BG and ZD. BX belongs to the highest social class, and ZD the lowest. This inference can be verified by combat effectiveness, as shown in Table 1. From the list of winning percentages, it can be inferred that the combat effectiveness (from strong to weak) among different families is: BX, SX, SQ, LD, G3, WX, WF, ZB, ST, TB, BG and ZD.

The experimental results are consistent with this ranking. There is a positive correlation among fighting power, occupied area and eating order. The stronger the family's combat effectiveness, the higher the eating order, and the larger the occupied area. The order and area of food reflect the fighting power to a certain extent, and also reflect the social class. The higher the social class, the higher the priority of resources.

Table 1. Combat effectiveness

	BX	SQ	WX	SX	G3	LD	WF	ZB	ST	BG	TB	ZD	Success rate
BX	*	0.83	0.83	0.17	1	0.58	1	1	1	1	1	1	9.41
SQ	0.17	*	0	0.5	1	0.5	1	1	1	1	0	0	6.17
WX	0.17	1	*	0	0.5	0.5	0.88	1	0	0	0	1	5.05
SX	0.83	0.25	0.89	*	0.33	1	0.5	0.57	0.67	1	1	1	8.04
G3	0	0.5	0.67	1	*	0	0	1	0	1	1	0	5.17
LD	0.33	0.5	0.5	0	0.67	*	0.67	1	0	1	1	0	5.67
WF	0	0	0.13	0.5	0.33	0.33	*	0.5	1	1	1	0	4.79
ZB	0	0	0	0.43	0.13	0	0.6	*	0.75	1	0	0	2.91
ST	0	0	0	0.33	0	0	0	0.25	*	1	0	0	1.58
BG	0	0	0	0	0	0	0	0	0	*	0	0	0
TB	0	0	0	0	0	0	0	0	0	1	*	0	1
ZD	0	0	0	0	0	0	0	0	0	0	0	*	0

5 Conclusion

In this paper, we develop a novel analysis framework based on fighting power of the golden snub-nosed monkey family and implicit feedback. We leverage tensor decomposition to reveal the potential relationship among the social class and food allocation. Fortunately, the results of tensor decomposition are consistent with the observations of animal protection experts.

Acknowledgments. This work was supported by the National Training Programs of Innovation and Entrepreneurship for Undergraduates under Grant. 201810697004, the Training Programs of Innovation and Entrepreneurship for Undergraduates of Shaanxi Province under Grant. 201807004, and the Training Programs of Innovation and Entrepreneurship for Undergraduates of Northwest University under Grant 342.

References

1. Wu, B.Q.: Patterns of spatial dispersion, locomotion and foraging behaviour in three groups of the Yunnan Snub-Nosed Langur (*Rhinopithecus bieti*). Folia Primatol (Basel) **60**(1–2), 63–71 (1993)
2. Maslow, A.H.: The role of dominance in the social and sexual behavior of infra-human primates: I. Observations at Vilas Park Zoo. J. Genet. Psychol. **48**(2), 261–277 (1936)
3. Kochhann, D.: Encyclopedia of animal cognition and behavior. In: Vonk, J., Shackelford, T. (eds.) Encyclopedia of Animal Cognition and Behavior, pp. 1–2. Springer, Cham (2017)
4. Wittemyer, G., Getz, W.M.: Hierarchical dominance structure and social organization in African Elephants, Loxodonta Africana. Anim. Behav. **73**(4), 671–681 (2007)
5. He, H., Zhao, H., Qi, X., Wang, X., Guo, S., Ji, W., Li, B.: Dominance rank of adult females and mating competition in Sichuan Snub-nosed Monkeys (Rhinopithecus roxellana) in the Qinling Mountains, China. Chin. Sci. Bull. **58**(18), 2205–2211 (2013)

6. Fleagle, J.G.: Evolutionary anthropology: issues, news, and reviews. In: Encyclopedia of Global Archaeology, p. 2706 (2014)
7. Chalmers, N.: Social Behavior in Primates. Edward Arnold, London (1979)
8. Wu, G., Wang, H.C., Fu, H.W., Zhao, J.Z., Yang, Y.Q.: Habitat selection of Guizhou Golden Monkey (Phinopithecus Roxellanae Brelichi) in Fanjing Mountain Biosphere Reserve, China. J. For. Res. **15**(3), 197–202 (2004)
9. Kolda, T.G., Bader, B.W.: Tensor decompositions and applications. SIAM Rev. **51**(3), 455–500 (2009)
10. Ailon, N., Chazelle, B.: Faster dimension reduction. Commun. ACM **53**(2), 97–104 (2010)
11. Yao, L., Sheng, Q.Z., Qin, Y., Wang, X., Shemshadi, A., He, Q.: Context-aware point-of-interest recommendation using tensor factorization with social regularization. In: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 1007–1010. ACM (2015)
12. Hao, F., Park, D.S., Yin, X., Wang, X., Phonexay, V.: A location-sensitive over-the-counter medicines recommender based on tensor decomposition. J. Supercomput., 1–18 (2018)
13. Tang, J., Shu, X., Qi, G.J., Li, Z., Wang, M., Yan, S., Jain, R.: Tri-clustered tensor completion for social-aware image tag refinement. IEEE Trans. Pattern Anal. Mach. Intell. **39**(8), 1662–1674 (2017)
14. Hu, W., Huang, J., Wang, Z., Wang, P., Yi, K., Wen, Y., Sun, L.: MUSA: Wi-Fi AP-assisted Video prefetching via tensor learning. In: 2017 IEEE/ACM 25th International Symposium on Quality of Service (IWQoS 2017), pp. 1–6. IEEE (2017)
15. Bhargava, P., Phan, T., Zhou, J., Lee, J.: Who, What, When, and Where: multi-dimensional collaborative recommendations using tensor factorization on sparse user-generated data. In: Proceedings of the 24th International Conference on World Wide Web, International World Wide Web Conferences Steering Committee, pp. 130–140 (2015)

SPARK-Based Partitioning Algorithm for k-Anonymization of Large RDFs

Odsuren Temuujin[1], Minhyuk Jeon[1], Kwangwon Seo[1], Jinhyun Ahn[2], and Dong-Hyuk Im[1(✉)]

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

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

Abstract. Privacy protection for resource description framework data is very important because RDF (i.e., linked data) is widely used in published data format in many areas, including government open data, health-care for individuals, and social relationships. As data can include private information belonging to individuals or companies and can make private information available to third parties, there are several anonymization models provided for preserving privacy in practice. k-anonymity has thus gained attention in research. Recently, several RDF anonymization models have been proposed. However, current approaches focus on a model and a metric for measuring information loss but do not consider large-scale RDF data. In this paper, we propose an efficient anonymizing method for large-scale RDF data. We develop a greedy partitioning algorithm (i.e., SPARK) for RDF anonymization. SPARK is a leading platform for big data processing. The results of experiments on synthetic datasets demonstrate that our proposed method requires less running time than previous methods.

Keywords: k-anonymity · Resource description framework · Apache SPARK · Data privacy

1 Introduction

The resource description framework (RDF) provides a metadata model to the web, which is a gigantic network full of data. Data accessed via the web is organized using RDF. Not unexpectedly, privacy protection issues have emerged with the increasing amount of data. As RDF is a widely used data format in various areas, such as government open data, individual health care, and social relationships, data can include private information belonging to individuals and companies, making private information available to third parties. Therefore, several protection techniques have been developed to ensure privacy. There are several anonymization models that have been used in practice, such as k-anonymity [1], l-diversity [2] and t-closeness [3]. The most well-known model for preserving privacy is k-anonymity, which provides privacy

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 292–298, 2020.
https://doi.org/10.1007/978-981-32-9244-4_41

protection by rendering data indistinguishable from at least k-1 others [1]. Therefore, several anonymization models for RDF data have been proposed [4, 5].

As existing anonymity approaches for RDF data aim to anonymize operations and metrics for precision and distortion of anonymized data, they do not provide a method to implement the kRDF anonymity model. Furthermore, they do not consider large-scale RDF data. Thus, we develop an anonymization algorithm for large-scale RDF data with a big data processing platform. To the best of our knowledge, ours is the first work devoted to apply partitioning algorithm for kRDF anonymization model in SPARK. The contribution of this paper is summarized as follows.

- We apply the existing partitioning algorithm to the RDF model.
- We provide a SPARK-based single-dimensional partitioning algorithm more efficiently than using platforms. SPARK [6] is the recently launched and leading platform for big data processing, extending the MapReduce model to more efficient types.
- Experiment results demonstrate the improved efficiency of our proposed approach compared to others.

2 The Proposed Approach

The proposed approach is based on a partitioning algorithm provided for anonymization. Our approach is a SPARK-based partitioning algorithm for anonymization of large RDF data.

The optimal k-anonymity is an NP hard problem [7]. Furthermore, the k-anonymity optimal partitioning algorithm is an NP hard problem. Therefore, we have numerous heuristic partitioning algorithms for k-anonymization [7, 8], and they often present effective results. Partitioning algorithms split the input dataset into two. Each dataset is then partitioned into two new partitions. This continues until each partition reaches a specified condition. Partitioning can be classified into single-dimensional and strict multidimensional partitioning.

2.1 Partitioning Algorithm for KR DF Anonymity in SPARK

We applied an anonymization model-based partitioning algorithm to the RDF model developed with SPARK. Our approach comprises preparation of data, data partitioning, and anonymization. Our algorithm is shown in Fig. 1.

First, we parse and set input data (i.e., quasi-identifier, sensitive identifiers..) from RDF data using the Apache Jena framework, is shown in Fig. 2. Apache Jena[1] is an open-source Java framework for building semantic web and linked-data applications. Then, we store data on resilient distributed dataset (RDD) in key and value pairs, representing a record in the input data. The RDD is simply a distributed collection of elements of SPARK's core abstraction.

[1] https://jena.apache.org/.

```
Phase 1: Transformation
    InputData: RDF data , k
    arrayList ← parse ( inputData )
    rdd ← createRDD ( arrayList )
    dimValues ← choose dimension( arrayList )
    key ← dimValues
    pairRdd ← map( key, rdd )
    medians ← isAllowableCut ( dimValues , k )
                  if ( dimValues.size <= k )
                      return null
                  else
                      splitVal ← find median( dimValues )
                      lhs ← { t ∈ dimValues : t. dimValue ≤ splitVal }
                      rhs ← { t ∈ dimValues : t. dimValue > splitVal}
                      if ( lhs.size >= k && rhs.size >= k )
                            isAllowableCut ( lhs, k )
                            isAllowableCut ( rhs, k )
              return splitVal
Phase 2: Action
    numOfParts = medians.count + 1
    partedRdd = customPartitioner ( pairRdd, numOfParts )
    Anonymize ( partedRdd )
```

Fig. 1. SPARK-based single-dimensional partitioning algorithm

SPARK's partitioning is available on all RDDs of key/value pairs and causes the system to group elements based on the function of each key. In our approach, the RDD split into *partitions* used by a custom partitioner, which we implemented. The number of tuples in the each part must be more than k, and the number of partitions depend on k. At the anonymization step, we use the mapPartitions SPARK operation to access all partitions and we use the generalization operation of tuples for each part to satisfy k.

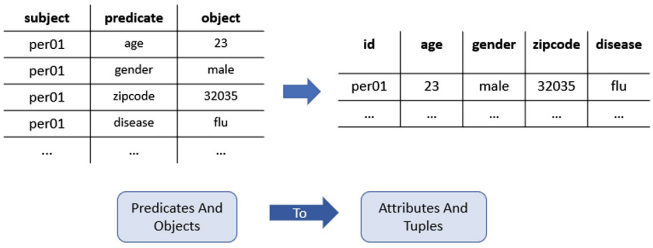


Fig. 2. Parse and setting input data

According to our proposed approach, an overview of SPARK-based partitioning for the 2-RDFanonymization process is shown in Fig. 3. Figure 3a shows the steps for creating and mapping RDDs, partitioning, and finding medians used for partitioning. One benefit of using SPARK is that you can write an application for parallel cluster executions in cluster mode. In cluster mode, SPARK uses one master coordinator (i.e., diver) and many distributed workers (i.e., executors). The driver is the process where the main method of the program runs. SPARK's executor processes run the individual tasks in each job. The schema of SPARK driver program and executors are shown in

Fig. 3a. We can see how cluster machines work with partitions from this schema. SPARK supports a variety of file systems, such as NFS, HDFS, and Amazon S3. We use HDFS to preserve anonymized data, as shown in Fig. 3b. The Hadoop distributed file system (HDFS) is a very popular and works well with SPARK.

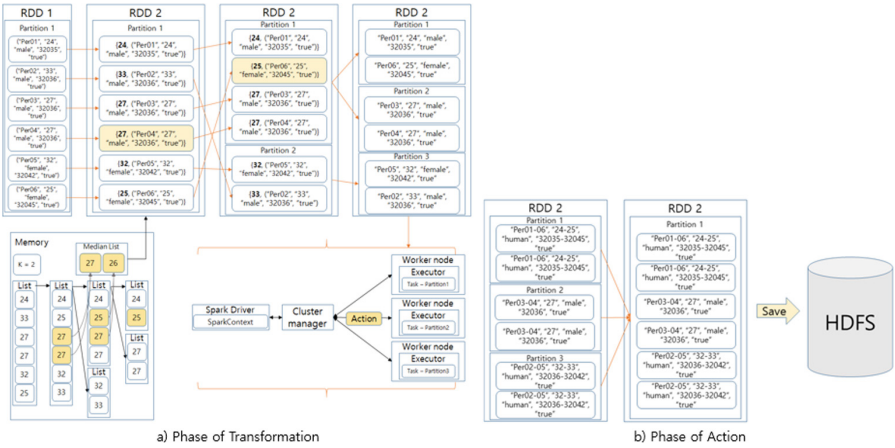


Fig. 3. Transformation and action phases

3 Experimental Evaluation

Our experiments evaluated anonymizations produced by partitioning algorithms to large-scale RDF data on SPARK and compared execution times on several cluster machines. Our cluster consists of five machines, each with a Xeon(R) 4-core CPU and 24-GB ram. Each cluster comprises a master and four workers. The master computer works as a driver, and we created one executor for each worker. Each executor uses three core CPUs with 3.10 GHz and 20-GB ram. For experiments, we used Lehigh University Benchmark (LUBM[2]) synthetic data, as shown in Table 1. From the LUBM dataset, we used only all-typed professor {*fullProfessor*, *associatedProfessor*, *assistantProfessor}* information for anonymization. Explicit identifiers of professors were removed from the dataset, and its predicates {*worksFor*, *researchInterest*, *bachelorDegreeFrom*, *masterDegreeFrom*, *doctorDegreeFrom*} were classified into quasi-identifiers (QID). An example of the LUMB dataset is shown in Fig. 4.

[2] http://swat.cse.lehigh.edu/projects/lubm/.

```
http://www.Department0.University0.edu/FullProfessor7

    ub:doctoralDegreeFrom        <http://www.University241.edu> ;
    ub:mastersDegreeFrom         <http://www.University901.edu> ;
    ub:researchInterest          "Research3" ;
    ub:undergraduateDegreeFrom   <http://www.University243.edu> ;
    ub:worksFor                  <http://www.Department0.University0.edu> .
```

Fig. 4. Example of LUBM dataset

Table 1. Synthetic data used in experiments

	LUBM 1	LUBM 10	LUBM 100	LUBM 1000
Size	6.4 Mb	82.2 Mb	876 Mb	8.62 Gb
Number of triples	≈100545	≈1272577	≈13405383	≈139415983

3.1 Performance Analysis

First, we compared single-dimensional in-memory data and SPARK-based single002Ddimensional partitioning algorithms using synthetic data. SPARK can run over a variety of cluster managers to access machines. In our case, we used SPARK standalone and yarn cluster managers. SPARK's standalone cluster manager comprised a master and multiple workers, each configured with memory and CPU cores. Yarn is a cluster manager introduced by Hadoop 2.0, installed on the same nodes as HDFS. The standalone cluster manager supports automatic recovery and Hadoop Yarn supports both of manual and automatic recovery through resource manager. Furthermore, yarn uses different encryption (i.e., Kerberos). Thus, it requires more processor time than the standalone cluster manager. Figure 5a shows the execution times of three different methods for the four datasets. In this experiment, the in-memory based approach works on a single machine. The SPARK (standalone) and SPARK (yarn) work on the cluster, consisting of one master and four workers. In this figure, the y-axis is expressed at a log scale and the x-axis is expressed by data size. For LUBM1000, the execution time of the in-memory-based approach is over the graph limitation. From this figure, we can see that the in-memory approach is useful for small data. However, for growing amount of data, our approach reduces execution time more than the in-memory approach. The duration of SPARK jobs on standalone cluster managers, using different cluster machines for different datasets, is shown in Fig. 5b. We used one master and four different workers. In most cases, the duration of SPARK job decreased while adding more clusters. Using more clusters is thus more efficient for job performance.

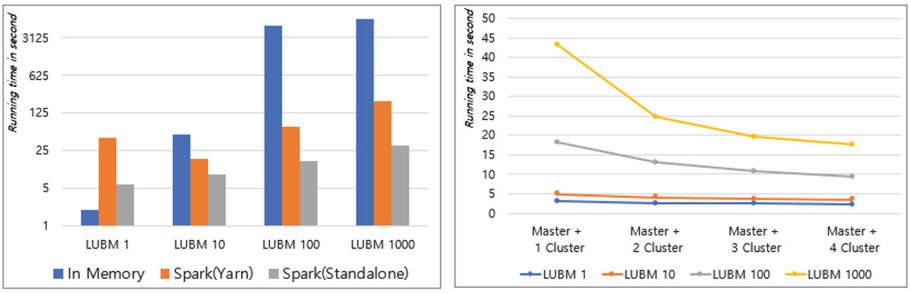


Fig. 5. Comparing performances

4 Conclusion and Future Work

In this paper, we introduced a SPARK-based k-anonymization partitioning algorithm for large RDF data. Our approach was based on a single-dimensional partitioning algorithm, which provided an anonymization model. The aim of the proposed approach was to develop a more efficient anonymization algorithm for large-scale RDF data, based on big data processing platform. An experimental evaluation demonstrated that our proposed approach was more efficient than in-memory-based approaches for anonymization. While growing data, our approach reduced the execution time more than previous approaches. Our future work will focus on other anonymization models such as l-diversity and t-closeness. Additionally, we plan to apply a strict multidimensional partitioning algorithm to SPARK.

Acknowledgement. This work was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2018-01417) supervised by the IITP (Institute for Information & Communications Technology Promotion) and IITP grant funded by the Korea government (MSIP) (No. R0113-15-0005, Development of a Unified Data Engineering Technology for Largescale Transaction Processing and Real-Time Complex Analytics) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. NRF-2018R1D1A1B07048380).

References

1. Sweeney, L.: k-anonymity: a model for protecting privacy. Int. J. Uncertainty Fuzziness Knowl.-Based Syst. **10**(05), 557–570 (2002)
2. Machanavajjhala, A., Gehrke, J., Kifer, D., Venkitasubramaniam, M.: l-diversity: privacy beyond k-anonymity. In: ICDE 2006, p. 24. IEEE (2006)
3. Radulovic, F., Garcia Castro, R., Gomez-Perez, A.: Towards the anonymization of RDF data (2015)
4. Heitmann, B., Hermsen, F., Decker, S.: k-RDF-neighbourhood anonymity: combining structural and attribute-based anonymization for linked data. In: PrivOn@ ISWC (2017)

5. Li, N., Li, T., Venkatasubramanian, S.: t-closeness: privacy beyond k-anonymity and l-diversity. In: IEEE 23rd International Conference on Data Engineering, ICDE 2007, pp. 106–115. IEEE (2007)
6. Zaharia, M., Xin, R.S., Wendell, P., Das, T., Armbrust, M., Dave, A., Meng, X., Rosen, J., Venkataraman, S., Franklin, M.J., et al.: Apache spark: a unified engine for big data processing. Commun. ACM **59**(11), 56–65 (2016)
7. LeFevre, K., DeWitt, D.J., Ramakrishnan, R.: Mondrian multidimensional k-anonymity. In: Proceedings of the 22nd International Conference on Data Engineering, ICDE 2006, p. 25. IEEE (2006)
8. Bayardo, R.J., Agrawal, R.: Data privacy through optimal k-anonymization. In: 21st International Conference on Data Engineering, ICDE 2005, Proceedings, pp. 217–228. IEEE (2005)

Toward a GUI-Based Comprehension of Software Architecture

Yeong-Seok Seo(✉)

Department of Computer Engineering, Yeungnam University, 280 Daehak-Ro, Gyeongsan, Gyeongbuk 38541, Republic of Korea
ysseo@yu.ac.kr

Abstract. Architecture-level software comprehension is important to maintain a software system. In particular, an architecture of a software system can assist software engineers in comprehending the high-level structure of the system. However, the knowledge of an architecture could be disappeared and incomprehensible over time, and this leads to poor sustainability. Therefore, it is necessary to recover the knowledge from the artifacts that are available even over time. To achieve this goal, we focus on Graphical User Interfaces (GUI) of a software system, which are mostly available in the life-time of the system. In this paper, we propose an approach to recover the knowledge of the system's architecture by using GUI. The case study is performed to evaluate the approach for an open-source software system and the results show that GUI provides useful evidences in identifying the high-level software structure.

Keywords: Architecture · Module view · Software comprehension · Graphical User Interface · Label

1 Introduction

Software engineers generally spend much of their time to comprehend unfamiliar software systems in the maintenance phase [1, 2]. To comprehend such software systems, software architecture play an important role by presenting the high-level of abstraction that provide a bridge between requirement and implementation [3]. In particular, the architecture of a software system abstracts its complex structure as more manageable and comprehensible high-level structure [4, 5]. As a result, the architecture let its software engineers know which part of it is in charge of what "responsibilities" to accomplish the requirements.

Despite of the importance of architecture-level software comprehension, the knowledge on a software architecture could disappear during software modification in the life-time of a software system by requirement changes, software failures, changes in personnel and so on. That could cause the loss of the memories or documents for the architecture or the inconsistencies between the documented and real software architecture. This leads to the inefficiency, incorrectness, or even impracticality of architecture-level software comprehension.

Software engineers, therefore, need to recover the knowledge of software architectures if there is no proper knowledge. Because the manual recovery requires heavy

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 299–304, 2020.
https://doi.org/10.1007/978-981-32-9244-4_42

labors, software engineers require the assistance of the techniques that provide the evidences to comprehend the architectures. The evidences provide a bridge between the available and comprehensible information sources in the maintenance phase and the elements of architectures. In the module-view, what "responsibilities" subsystems provide to the rest of a software system is important information to comprehend an architecture. Therefore, the evidences provide a bridge between information sources and the responsibilities of subsystems.

In this study, we focus on Graphical User Interfaces (GUIs), which interface with users through a graphical screen display having text embedded in it, to provide the evidences to comprehend software architectures, especially their module-view. GUIs are useful information sources for that purpose because the interfacial role of GUIs makes them not only always available to the users of software systems at least after software systems are released, but also maintained to be user-comprehensible. Therefore, we propose a GUI-based labeling approach to provide a bridge between the user-comprehensible information from GUIs and the responsibilities of subsystems.

The remainder of this paper is organized as follows: Sect. 2 introduces the proposed approach. Section 3 shows a case study to evaluate the proposed approach, and Sect. 4 concludes this paper and suggests future work.

2 GUI-Based Labeling for Architecture Comprehension

Based on the GUI properties, we propose a GUI-based labeling to support understanding of a software architecture.

2.1 Basic Concept of GUI-Based Labeling

There are several following properties of GUI and software architecture.

- The architecture of a software system consists of subsystems that provide their own responsibilities to the rest of the system.
- The functionalities of a software system are presented to users through the components on its GUI, such as buttons, panels and windows. One of the very natures of GUIs is that the functionalities (components) are presented (deployed) on GUIs in the hierarchical manner [6] (e.g., the hierarchical structure in Fig. 1-(c) of the GUI in Fig. 1-(a) and (b)). The hierarchical structure of GUIs is built by using two types of components: *actionable* and *non-actionable*. Actionable components (e.g., buttons) are the access points on GUIs to the functionalities of software systems. Non-actionable components (e.g., panels and windows) are the frames to organize the actionable components on GUIs in the hierarchical manner.
- Users can comprehend the functionalities of components through the graphical information that users take from the GUIs. For example, users can comprehend the functionalities of *Blur* in Fig. 1-(a) through the graphical information in Fig. 1-(b) (In the rest of this study, the term "component" includes not only a logical component on a GUI, but also the graphical information about it).

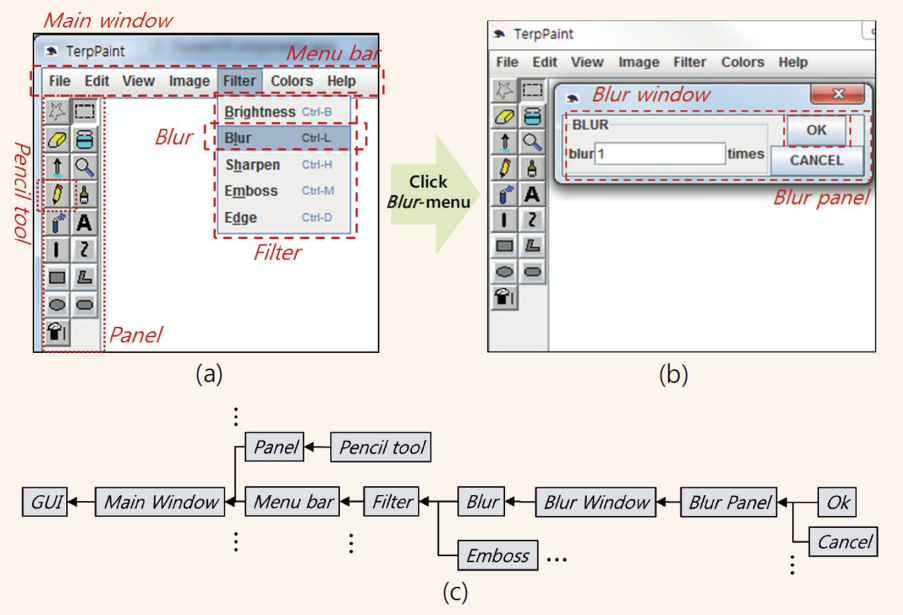


Fig. 1. Example of hierarchical structures of GUIs

- When users interact with actionable components, they make subsystems executed to perform their functionalities.

The goal of our approach is to provide the evidences for *Responsibility of Subsystem* by "bridging" it to user-comprehensible *Functionality of Component*. The bridge can be built by analyzing the relationships between *Actionable* and *Subsystem*. Our approach analyzes the hierarchical structure of GUI to extend the relationships to *Non-actionable*. As a result, our approach provides *Functionality of Components* to comprehend *Responsibility of Subsystem* in terms of functionalities.

2.2 Bridging Between Subsystem and Component

GUI-based labeling builds the bridge between subsystems and components through the three phases: (1) GUI analysis, (2) trace analysis, and (3) bridge construction.

- Phase 1: GUI Analysis. In this phase, the hierarchical structure of a GUI is extracted into a component hierarchy. A component hierarchy is built based on the two types of hierarchical relationships between the components on the GUIs: *laid-on* and *popped-up* relationships. A laid-on relationship is the structural relationship between two components, one of which is laid on another. For example, in Fig. 1-(a), there is a laid-on relationship between node *Filter* and *Blur*. In the functional point of view, *Blur* is a sub-functionality of *Filter*. A popped-up relationship is the relationship between two components, one of which is popped up by activating another. For example, there is a popped-up relationship between *Blur* in Fig. 1-(a)

and *Blur window* in Fig. 1-(b). In the functional point of view, the window provides the (sub-)functionalities related to blurring. Therefore, a component hierarchy presents functional decomposition of a software system.

- Phase 2: Trace Analysis. This phase builds the bridge between *Actionable* and *Subsystem* by tracing the subsystems that are executed when each actionable component is activated. As the result of this phase, the logic coverage Info. is collected, which indicates which part of a subsystem, such as lines, functions and classes in the subsystem, are executed by an actionable component. When $\boldsymbol{S}$ is the set of subsystems in a software system, a logic coverage Information of an actionable component, $c_{act} \in \boldsymbol{C}$, is defined as follows: $(S_i, m_j) \in LC(c_{act})$, where $S_i \in \boldsymbol{S}$: a subsystem traced by activating c_{act}, $m_j \in S_i$: a traced constituent of S_i by activating c_{act}
- Phase 3: Bridge Construction. This phase completes the bridge between components in a component hierarchy and subsystems based on the results of phase 1 and 2 (i.e. the component hierarchy and logic coverage Information). As the result of bridge construction, a subsystem-by-component matrix is built. Each cell of the matrix presents how much part of a subsystem, $\boldsymbol{S}$, are covered by a component, c.

2.3 Subsystem Comprehension Through GUI-Based Labels

The labels are defined as the components whose coverage is larger than a given threshold α for a subsystem in the subsystem-by-component matrix. For example, $S1$ has two labels: *Filter* and *Menu* with 80% coverage. Software engineers can and the corresponding components of the labels on the GUI. The graphical information of *Menu* and *Filter* from the GUI let the software engineers comprehend the functionalities of $S1$ easily (i.e. filtering a figure on the canvas). Through the same process for the other subsystems, software engineers could comprehend the functional responsibilities of the other subsystems and the architecture of the software system in the functional point of view.

3 Case Study

To examine that our approach can produce acceptable and useful labels for a given architecture of a software system, we investigate the following question: Can the structure of the GUI really provide the functional decomposition of the subsystem?

We apply our approach to an open-source software system to answer the questions. As the subject system, we choose Weka that is a software system written by Java for various data mining tasks [7]. Weka has a ready-made architecture that has a hierarchical structure (e.g., subsystem weka.gui contains weka.gui.experiment). The hierarchy consists of 76 subsystems. There are 1,049 classes in the subsystems.

In order to answer the question, we present the component hierarchy of Weka to analyze that it presents meaningful functional decomposition of Weka. To create the component hierarchy, Weka has been executed overnight (i.e. about 3,000 interactions on Weka's GUI). As a result, a component hierarchy with 333 nodes is built. We refine the hierarchy to reduce meaningless nodes in terms of the functional decomposition

(e.g., a node with a single child). As a result, our prototype extracts a component hierarchy with 205 nodes.

The component hierarchy of Weka is presented in Fig. 2. Due to the limitation of space, we only present the partial component hierarchy. In particular, we present the high-level structure of the hierarchy because the hierarchy presents more meaningful functional decomposition of Weka in the high-level than the low-level.

Figure 2 and Table 1 show that the component hierarchy presents the functionalities of Weka. Figure 3 present an example of the label for $n_{1\text{-}3\text{-}2}$ in Fig. 2. The functionalities of $n_{1\text{-}3\text{-}2}$ in Fig. 2 can be identified through the snapshot in Fig. 3: choosing, testing and running classifier. The component hierarchy comes from the menu-hierarchy and tab-structure on Weka's GUI. Therefore, we can conclude that the structure of the GUI can provide the functional decomposition of Weka.

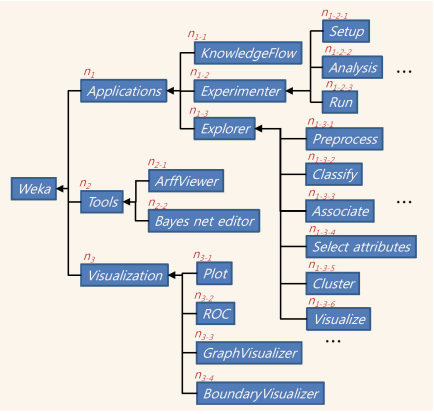


Fig. 2. The (partial) component hierarchy of Weka

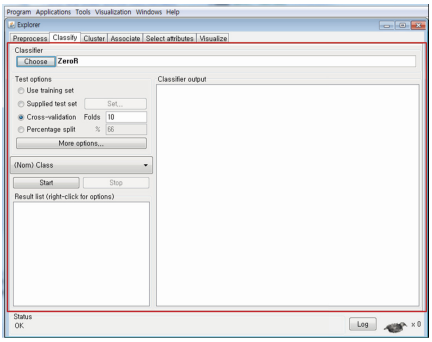


Fig. 3. Example of the label for $n_{1\text{-}3\text{-}2}$ in Table 1

Table 1. The functionalities of the labels in Fig. 2

n_{ID}	Functionality
n_1	Apply data mining algorithms to data
$n_{1\text{-}1}$	Explore data as a graphical front end
$n_{1\text{-}2}$	Create, run, modify and analyze experiments
$n_{1\text{-}2\text{-}1}$	Set parameters for an experiment
$n_{1\text{-}2\text{-}2}$	Analyze the experiment
$n_{1\text{-}2\text{-}3}$	Run the experiment
$n_{1\text{-}3}$	Explore data
n_2	Manipulating data or graphs
$n_{2\text{-}1}$	View and edit.ar data
$n_{2\text{-}2}$	View and edit graphs
n_3	Visualize various types of data

4 Conclusion

The functional decomposition from GUIs could be used in various future studies of software re-engineering: architecture re-construction, bad smells detection and so on. It is a promising future work to conduct case studies with various software systems and figure out more issues to be considered for applying our approach in practice. It is worth to study of labeling for various kinds of responsibilities as well as the functional responsibilities of subsystems. Our approach will also be able to extend to cover other views of software architectures. We expect that more useful information can be extracted from GUIs to use the information for architecture comprehension.

Acknowledgment. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2017R1C1B5018295).

References

1. Ko, A., Myers, B., Coblenz, M., Aung, H.: An exploratory study of how developers seek, relate, and collect relevant information during software maintenance tasks. IEEE Trans. Softw. Eng. **32**, 971–987 (2006)
2. Fernandez-Saez, A.M., Chaudron, M.R.V., Genero, M.: An industrial case study on the use of UML in software maintenance and its perceived benefits and hurdles. Empir. Softw. Eng. **23**, 3281–3345 (2018)
3. Ducasse, S., Pollet, D.: Software architecture reconstruction: a process-oriented taxonomy. IEEE Trans. Softw. Eng. **35**, 573–591 (2009)
4. Maqbool, O., Babri, H.: Hierarchical clustering for software architecture recovery. IEEE Trans. Softw. Eng. **33**, 759–780 (2007)
5. Lutellier, T., Chollak, D., Garcia, J., Tan, L., Rayside, D., Medvidovic, N., Kroeger, R.: Measuring the impact of code dependencies on software architecture recovery techniques. IEEE Trans. Softw. Eng. **44**, 159–181 (2018)
6. Memon, A.M.: A comprehensive framework for testing graphical user interfaces. Ph.D. Dissertation (2001)
7. Weka. https://www.cs.waikato.ac.nz/ml/weka/

UAV Anomaly Detection with Distributed Artificial Intelligence Based on LSTM-AE and AE

Gimin Bae and Inwhee Joe(✉)

Department of Computer and Software, Hanyang University,
Wangsimni-ro, 222, Seoul 04763, Korea
{baegimin, iwjoe}@hanyang.ac.kr

Abstract. In this paper, we propose a novel method for UAV anomaly detection in the distributed artificial intelligence environment by using deep learning models. In the conventional artificial intelligence environment, a lot of computing power is required for anomaly detection, so it is not suitable to the UAV environment based on embedded systems. For UAV anomaly detection, distributed artificial intelligence with DPS (Distributed Problem Solving) and MAS (Multi-Agent System) is applied using LSTM-AE and AE models. The experimental results show that the proposed method performs well for anomaly detection in the UAV environment.

Keywords: UAV · Anomaly detection · Intrusion detection · Scoring · LSTM-AE

1 Introduction

Unmanned Aerial Vehicle (UAV) has been developed for military use but recently it has been rapidly spreading in various fields such as logistics transportation, agriculture, information communication, disaster observation, environment, transportation, personal hobby, etc. The existing UAV is simply a small airplane, but recently it has created a number of new business models by converging with the fields of Internet communication, agriculture and environmental protection (Table 1).

Along with the proliferation of UAVs, threats targeting UAVs are also increasing. If it is used for mechanical defects or terrorism of UAV, it will cause enormous damage. Therefore, to protect the UAV from potential threats to UAV security threats must be detected and prevented in advance.

In this paper, we propose an anomaly detection method based on distributed artificial intelligence using LSTM-AE (Long Short Term Memory Network Auto-Encoder) to detect and prevent security threats using flight data generated from UAV.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 305–310, 2020.
https://doi.org/10.1007/978-981-32-9244-4_43

Table 1. UAV security vulnerability

Vulnerability	Description
GPS spoofing	It is a method to manipulate the position and time so that the receiver of the UAV selects the signal sent by the GPS Spoofing attacker without selecting the signal received from the satellite
GPS jamming	How to send a strong disturbance signal at the same frequency as GPS to prevent normal signal reception
Destruction	Destruction of drones due to external attack
Malware	Exposing UAV-specific malicious code threats during UAV software update or external interface connection
Tapping	Sensitive information leakage collected by drones through radio eavesdropping
Terror	Terrorism of Major Facilities Using Drones

2 Related Work and Background

2.1 Related Work

Liu and Ding proposal of real-time Anomaly Detection of UAV Flight Data Using K Nearest Neighbor Search (KNNS) Algorithm [1]. Yong, Duan, et al. proposal of detection method of abnormal number of UAV sensor data through kernel principal component analysis (KPCA) [2]. Jeong et al. proposal of real-time anomaly detection system for uav control system [3]. Lu et al. proposal of an anomaly detection system to prevent the motor of the UAV from operating at abnormal temperatures [4]. He et al. proposal of an online algorithm for detecting and mitigating flight data deformation using anomaly detection and mitigation algorithm based on online subspace tracking [5]. Choudhary et al. in a networked UAV environment existing intrusion detection system research and UAV intrusion detection system proposal [6].

2.2 LSTM

As a proposed LSTM cell in 1997, it not only solves the long-term dependence problem of RNN cells, but also learns quickly. The structure of the LSTM cell is shown in Fig. 1 [7]. In the LSTM cell, the state is divided into two vectors. H is a short-term state, and c_t is a long-term state. At the core of LSTM Cell when the network is in the long-term state (c_t) it learns what to remember, what to delete, and what to read. Long-term memory passes through the cell from left to right, passing a forget-gate, losing some memory (information), and then adding a new memory part from the input gate with the addition (+) operation. The generated c_t is output immediately without any additional operation, and this long-term memory c_t goes through a process of deleting and adding some memories at each time step. After addition, c_t is copied and transferred to the function of the output gate to produce the short-term state h_t and the output y_t of the cell. The three layers that output f_t, i_t, and o_t use sigmoid (logistic) as the activation function. The output of the sigmoid function is 0 to 1, and this output value is input to the element-wise multiplication operation of each forget, input, and output

gates. Therefore, if the output is 0, the gate is closed. If it is 1, the gate is opened. Therefore, f_t, i_t, and o_t are called gate controllers.

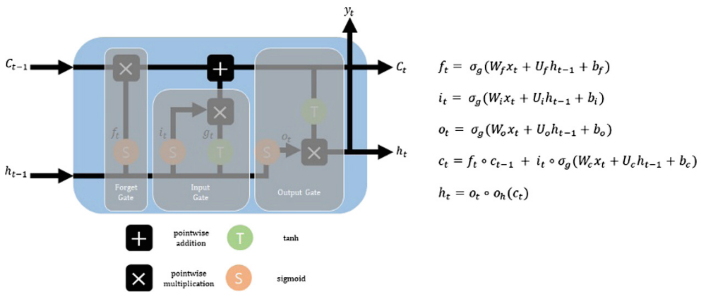


Fig. 1. LSTM cell

2.3 Distributed Artificial Intelligence

DAI (Distributed Artificial Intelligence) is a technique for distributing and resolving intelligent agents with complex problems that are difficult to solve in one place. Local optimization is performed through local information, and these results are solved in the form of generating global optimization effects [8]. DAI is divided into DSP (Distributed Problem Solving), MAS (Multi-Agent System), SI (Swarm Intelligence) and PAI (Parallel AI). A DSP is a way to divide the problem for one big problem, process it in a distributed node, and mash up the partial results to get the final result. In a heterogeneous distributed node environment with different interests and objectives, each node derives its results from the process of collaboration, competition, consent, and denial.

3 Proposed Method

In this paper proposes UAV anomaly detection method using LSTM-AE [9] model based on distributed artificial intelligence to improve safety and security of UAV by using flight data and external environmental information collected from UAV.

In the case of the proposed UAV anomaly detection, proceed to step 2. UAV uses UAV status information to perform UAV anomaly detection (Low-Level). In the GCC (Ground Control Center), the overall anomaly detection (High-Level) of the mission is performed using the UAV status information, the position information, the speed information, and the environmental information. through the anomaly detection result the data restoration value and the next sequence corresponding to the current sequence are predicted, and the restored value and the predicted value are judged as normal or abnormal based on the threshold value.

Figure 2(a) is LSTM-AE and Fig. 2(b) is AutoEncoder used in this paper. For time series data, use LSTM-AE model. For other data, use AutoEncoder model. In case of

LSTM-AE, it outputs two outputs (Reconstruction Error value and predicted value) and outputs one output (Reconstruction Error value) for AutoEncoder.

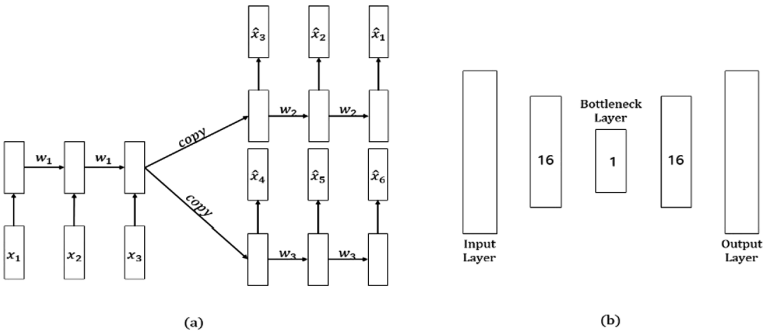


Fig. 2. (a) LSTM-AE & (b) AutoEncoder

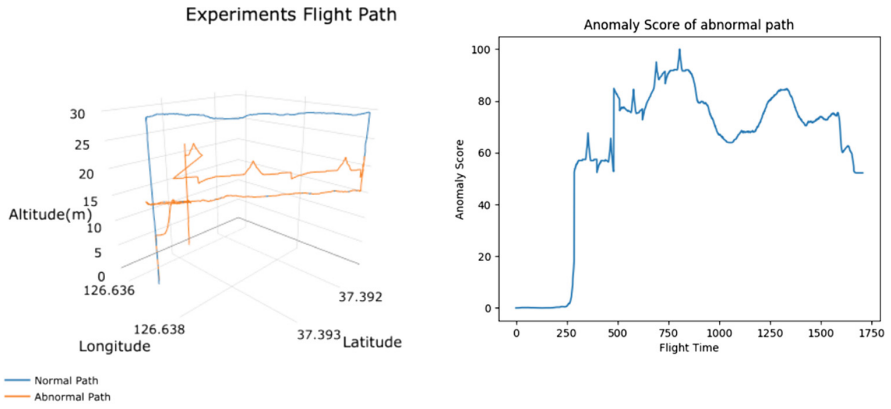
4 Experiments

The experiment in this paper utilizes UAV flight data from DJI Inspire 2, simulator and raspberry pi base UAV. The UAV flight path is predefined and assumed to be a normal mission. In the case of abnormal flight data, it is defined as an undefined flight situation. In case of anomaly detection (low-level) executed in UAV, pre-learned model of UAV is load and anomaly detection is performed using UAV flight data generated at the time of flight. For anomaly detection (high-level) that runs on GCC, it uses Cloud Server provided by Naver Cloud Platform. It is also assumed that cloud server receives data from UAV through network socket communication.

Figure 3(a) is the normal flight path and Fig. 3(b) is the abnormal flight path anomaly score. Anomaly score was calculated using mMAPE [10].

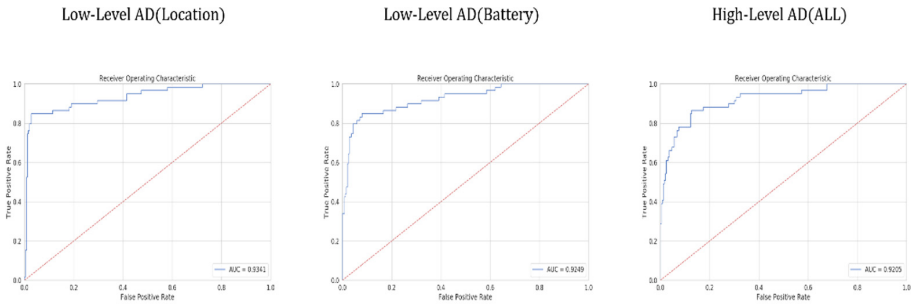
Figure 4(a) is the low-level anomaly detection result and Fig. 4(b) is the high-level anomaly detection result. For low-level anomaly detection, used UAV location and battery data. For high-level anomaly detection, used data corresponding to UAV position, battery, attitude, speed, and external environment.

Figure 5 shows the learning time the conventional anomaly detection method and proposed anomaly detection method in Raspberry Pi based UAV. There are two methods proposed in this paper: DAI-MAS with applied MAS and DAI with non-applied MAS Using the conventional anomaly detection method, it takes 30 s of learning time per epoch, and it is not suitable to apply to UAV. In the proposed method, DAI consumes 5 s per epoch and DAI-MAS consumes 1.2 s per epoch. This ensures that the proposed method is better suited to the UAV environment than the conventional approach.



(a) (b)

Fig. 3. Flight path & anomaly score



(a) (b)

Fig. 4. Low-level AD & High-level AD result (ROC curve)

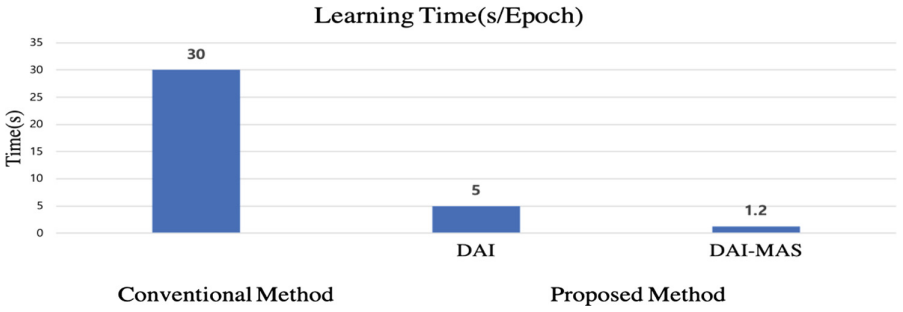


Fig. 5. Device performance comparison

5 Conclusion

Unmanned aerial vehicles (UAVs) have recently been developed for military use but are rapidly spreading in various fields such as logistics, agriculture, telecommunications, disaster observation, environment, transportation, and hobby. Along with the proliferation of UAVs, threats targeting UAVs are also increasing. If it is used for mechanical defects or terrorism of UAV, it will cause enormous damage. Therefore, UAVs should be detected and prevented in advance to protect them from potential threats. However, the conventional anomaly detection method is not suitable for UAV environment because it requires a lot of computing performance. In this paper, we propose a method to detect abnormal symptoms in distributed artificial intelligence environment using LSTM-AE and AE models. To this end, UAV anomaly detection was experimented with distributed artificial placement (distributed problem solving) and MAS (Multi Agent System). Experimental results show that the proposed method is suitable for performing anomaly detection in UAV environment.

Acknowledgment. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science, ICT & Future Planning) (NO. 2016R1A2B4013118).

References

1. Liu, Y., Ding, W.: A KNNS based anomaly detection method applied for UAV flight data stream. In: Prognostics and System Health Management Conference (PHM). IEEE (2015)
2. Yong, D., et al.: Unmanned aerial vehicle sensor data anomaly detection using kernel principle component analysis. In: 2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI). IEEE (2017)
3. Jeong, H.-J., et al.: Machine learning-based real-time anomaly detection for unmanned aerial vehicles with a cloud server. J. Internet Technol. **18**(4), 823–832 (2017)
4. Lu, H., et al.: Motor anomaly detection for unmanned aerial vehicles using reinforcement learning. IEEE Internet Things J. **5**(4), 2315–2322 (2018)
5. He, Y., et al.: ADMOST: UAV flight data anomaly detection and mitigation via online subspace tracking. IEEE Trans. Instrum. Meas. **99**, 1–10 (2018)
6. Choudhary, G., et al.: Intrusion detection systems for networked unmanned aerial vehicles: a survey. In: 2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC). IEEE (2018)
7. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997)
8. Park, J.H., et al.: IoT based distributed intelligence technology for hyper-connected space. Electron. Telecommun. Trends **33**(1), 11–19 (2018)
9. Srivastava, N., Mansimov, E., Salakhudinov, R.: Unsupervised learning of video representations using LSTMs. In: International Conference on Machine Learning (2015)
10. Shin, K.-H.: Estimation method of predicted time series data based on absolute maximum value. J. Energy Eng. **27**(4), 103 (2018)

Data-Request-Packet Broadcasting Approach for Mobile Opportunistic Networks with Unreliable Links

Lichen Zhang[1,2], Xiaochun Zhang[2], Sui Yu[2], Longjiang Guo[1,2(✉)], Peng Li[1,2], Meirui Ren[1,2], and Xiaoming Wang[1,2]

[1] Engineering Laboratory of Teaching Information Technology of Shaanxi Province, Xi'an 710119, China
[2] School of Computer Science, Shaanxi Normal University, Xi'an 710119, China
{zhanglichen, xiaochuan_zhang, catfish_sui, longjiangguo, lipeng, meiruiren, wangxm}@snnu.edu.cn

Abstract. With the wide application of mobile smart devices carried by humans, content dissemination by using Device-to-Device wireless communication technologies in Mobile Opportunistic Networks (MONs) attracts numerous attentions. However, traditional content dissemination approaches have not considered the efficiency of broadcasting interest packets to the providers in MONs with unreliable links. To increase the efficiency of broadcasting interest packets, we propose two data-request-packets broadcasting approaches with the goal of decreasing energy and communication consumption, in which the policy of self-adapting broadcasting time interval is applied. The extensive simulation results show that our proposed approaches greatly prolong the network lifetime, while maintaining higher delivery ratios compared with traditional approach.

Keywords: Broadcast · Data-request · Mobile opportunistic network

1 Introduction

With the development of short-distance wireless communication technologies and mobile smart devices (e.g., smartphones, smart glasses carried by humans), mobile opportunistic networks (MONs) have been widely applied on various fields such as intelligent transportation, intelligent healthcare, and business applications [1]. In MONs, nodes can directly communicate with D2D communication technology when two nodes are within the range of their wireless communication range, and then needs no communication infrastructure, which makes it especially suitable for those infrastructure-less environments, in which humans could apply convenient content sharing based on the D2D [2]. Therefore, MONs have been attracted numerous attentions in the topic of routing [3–6], content dissemination [7–11], of which data distribution (also called content dissemination) is one of the key topics.

Data distribution [5] in MONs is a service provider would like to disseminate some contents to those nodes (also called consumers) who are interested in the contents.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 311–317, 2020.
https://doi.org/10.1007/978-981-32-9244-4_44

Traditional approaches usually assume that the providers know in advance all those consumer nodes or apply the flooding-based methods to broadcast the interest packets (called data-request-packets) to the providers. In fact, the flooding-based approach will consume great energy and communication resources even though the size of data-request-packets is small, which is not suitable for energy and communication restricted devices in MONs [2]. Therefore, it is meaningful and important to present novel data-request-packets broadcasting approaches to decrease the resource consumption and then prolong the network lifetime.

Currently, a variety of approaches have been proposed for routing, content sharing and data transmission [3–11]. However, the efficiency remains a big problem due to the reason that data-request-packets are usually small and the traditional approaches are designed for the large size of data packets. Moreover, the consumption of energy and communication in broadcasting small packets are ignored in most approaches, which should be considered in data-request-packet broadcasting approaches.

Motivated by the above observations, in this paper, the traditional flooding-based approaches and two improved ones with the goal of decreasing the consumption of energy and communication resources have been provided. First, the unreliable links in MONs are considered in data-request-packet broadcasting. Second, the numbers of broadcasting each packet and connection with other nodes are controlled with the goal of decreasing energy and communication consumption while maintaining a slightly higher network performance. Third, the policy of self-adapting broadcasting time interval is applied, in which the time interval of broadcasting a data-request-packet is exponentially increased with the number of relays to other nodes.

2 Network Model and Assumptions

A mobile opportunistic network consists of a variety of mobile nodes with limited energy and storage resources randomly move around thus leading to a number of isolated sub-networks. When some nodes move within the range of wireless transmission, wireless communication and message exchange could occur.

Suppose there is only one broadcasting channel used for the meta-data exchange and the single channel is unreliable which means the data transmission may disrupt frequently. We use IEEE 802.11 protocol in the connection establishment process between two nodes to avoid the hidden terminal problem. In MONs, each node will consume energy in the process of listening channels, sending and receiving packets. Suppose that a data-request-packet is a 4-tuple, represented by $m = (sid, cid, t, ttl)$, in which $m.sid$ denotes the node generating the packet m, $m.cid$ denotes the content id of packet m, $m.t$ denotes the time when packet m is generated, and $m.ttl$ denotes the remaining time to live. Note that, $m.cid$ usually means the interest name or the logical address of content.

To make things easier to illustrate, we assume there exists only one node (usually called provider) for each different data-request-packet. Therefore, the basic functional goal of this paper is provide efficient approaches of broadcasting all data-request-packets over single channel to their related providers with the minimum consumption of energy and storage resources.

3 Efficient Data-Request-Packet Broadcasting Approaches

3.1 Traditional Approach

In the traditional approach, described in Table 1, each node maintains a list of received data-request-packets in its memory, and periodically broadcasts id information to let other nodes build connections. Whenever a connection is built, the list of data-request-packets of both nodes are exchanged and updated.

Table 1. The traditional data-request-packet broadcasting algorithm.

```
//Each node (denoted by v) listens the channel
1. every α time units node v broadcasts its own id if the channel is idle, or
   waits a random time units to monitor the channel again.
2. if node v hears a message from another node(denoted by v') then
3.     node v tries to establish a connection to v_j by using 802.11 protocol.
4.     if connection is established then
5.         nodes v and v' exchange their data-request-packets by using 802.11
             protocol, and drop the packets that have been stored.
6.         close connection.
7.         endif
8. endif
```

Each node maintains a local timer with the initial value of α and decreases with time. A node broadcasts if the clock reaches 0, and sets the timer to α again when the broadcasting process ends. The local timer of a node pauses when the node is trying to connect with other nodes, and starts when the connection breaks.

3.2 An Improved Broadcasting Approach (Improved-1)

To decrease unnecessary data transmission, each node v maintains a list (denoted by L_v) of recently connected nodes. Each element in list L_v is a 2-tuple, denoted by $e = (id, ttl)$, where $e.id$ is the id of node v with which node v has been connected, and $e.ttl$ is the time to live of element e. The initial value of $e.ttl$ is set to T. The main difference between the Improved-1 algorithm and the Traditional Approach is that a node will not built a connection to those nodes that have been connected within T time units, which leads to great decrease of the number of unnecessary connections.

Note that, $e.ttl$ decreases with time, and if reaching to 0, the element e will be dropped. In the Improved-1 Algorithm, because of the records of recently connected nodes, each node will avoid a huge number of connections with its neighbors, with which data-request-packets have been exchanged.

3.3 An Improved Broadcasting Approach (Improved-2)

In the Improved-2 algorithm, described in Table 2, each data-request-packet contains two new information, that is k and r, which are the number of relays of the packet and the remaining time for next relay. The value of k is 0 for the newly generated data-request-packet, and increased by 1 for each relay. Whenever a packet is relayed, the value of r is set to 2^k time units, and decreases with time. When r equals to 0, the related packet needs to be broadcast again. Each node maintains a list of ids of data-request-packets that needs to be broadcast. Whenever a node hears a broadcast packet from its neighbor, the node checks whether there exists an unreceived packet. If so, the node tries to build connection, otherwise, remains silent.

Table 2. The Improved broadcasting algorithm (Improved-2).

1.	for all those data-request-packets whose r equals to 0, node v broadcasts those ids with its own id if the channel is idle, or waits a random time units to monitor the channel again.
2.	**if** node v hears a broadcast message from another node (denoted by v') **then**
3.	**if** there exist a data-request-packet whose id is in the broadcast message **then**
4.	node v tries to establish a connection to v_j via 802.11protocol.
5.	**if** connection is established **then**
6.	nodes v and v' exchange those data-request-packets whose $r = 0$ by using 802.11 protocol. For each exchanged packet, increase k by 1 and set new value of r to 2^k.
7.	node v removes those ids whose related packets have been sent to v' in the broadcast message.
8.	close connection.
9.	**endif**
10.	**endif**
11.	**endif**

In Improved-2 broadcasting algorithm, the number of relays of each data-request-packet is roughly controlled. For every relay of a packet, the waiting time for the next relay will be doubled. Thus, in the case that there exist a great amount of data-request-packets, both the number of broadcasting and connections are greatly decreased, leading to decrease the number of wireless interferences, and also save great energy and communication resources. For a data-request-packet m, the maximum number of relays by a given node is $\log_2 N$, where N is the TTL of packet m.

4 Simulation and Results

4.1 Simulation Settings

A number of nodes are placed randomly and move continuously according to the RWP (Random Way Point) mobility model with a speed of 0.5−1 m/s with pause time 1 −5 s, in which each node has a transmission range of 80 m. The buffer size of a node is set to 1 MB, since the size of data-request-packet is usually very small. The transmission speed of each node is 80 KB/s, the number of nodes ranges from 200 to 1000, and the side length of the area changes from 200 m to 600 m. The generation rate of a packet is set to 100−200 s. The simulation time is 20000 s, with no warming time. Besides, we set the wireless loss rate as 0.2 and the node's buffer size is 1M.

4.2 Simulation Results

(1) Impact of Different Network Sizes. In Fig. 1, the network size ranges from 200 to 1000. The results show that the average flooding delay and energy cost increase as the network size increases accordingly. This is because as the number of nodes increases, nodes will encounter more frequently and consume more energy in exchanging data packets, so that our improved approach could effectively reduce unnecessary connections and delay by considering reduce reconnections in certain period and prolong network lifetime. Figure 1(c) shows that the impact of data packet delivery ratio with the changes of network sizes. We can see that our approach could achieve higher delivery ratio in a greater network size.

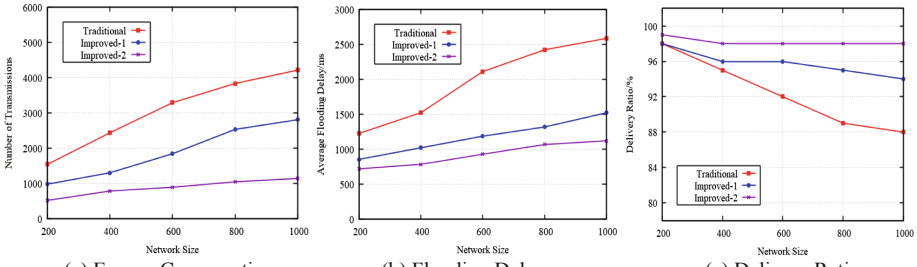
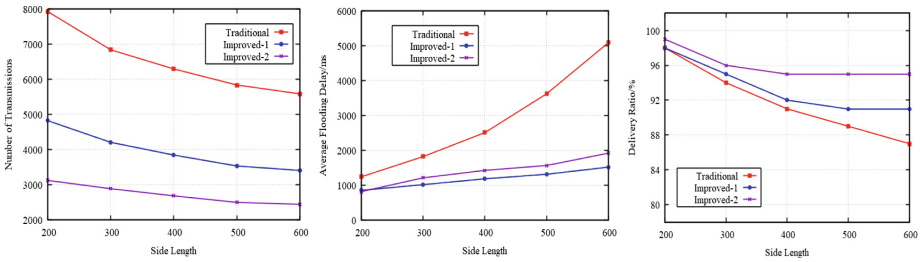


Fig. 1. Flooding performance in networks with different network size.

(2) Impact of Different Network Densities. For different network densities, 600 nodes are deployed in an area. As shown in Fig. 2, energy cost goes down as the network density decreases. Moreover, the Improved-2 performs much better than traditional and Improved-1. This is because as the diameter of the network becomes larger, nodes will record the recent connection information and broadcast stored packets after certain sleep time to reduce the delay and save energy.



(a) Energy Consumption (b) Flooding Delay (c) Delivery Ratio

Fig. 2. Flooding performance in networks with different network densities.

5 Conclusions

In order to save energy and communication consumption in packet requesting in MONs with the existence of unreliable links, novel approaches have been provided, where a policy is deeply applied to control the number of broadcasting and connection establishment. The results show the improved network lifetime. Some future works focus on the improvement of other research topics in MONs with unreliable links with the consideration of wireless broadcast nature and the consumption of energy and communication cause by small packets.

Acknowledgements. This research is partially supported by the National Key R&D Program of China (No. 2017YFB1402102), the NSF of China (No. 61877037), the Natural Science Basis Research Plan in Shaanxi Province of China (No. 2017JM6060), and the Fundamental Research Funds for the Central Universities of China (Nos. GK201903090).

References

1. Zhao, Y., Song, W.: Survey on social-aware data dissemination over mobile wireless networks. IEEE Access **5**, 6049–6059 (2017)
2. Trifunovic, S., Kouyoumdjieva, S., Dist, B., Pajevic, L., Karsson, G., Plattner, B.: A decade of research in opportunistic networks: challenges, relevance, and future directions. IEEE Commun. Mag. **55**, 168–173 (2017)
3. Zhang, L., Cai, Z., Lu, J., Wang, X.: Spacial mobility prediction based routing scheme in delay/disruption-tolerant networks. In: Proceedings of International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI 2014), pp. 274–279. IEEE Press, New York (2014)
4. Li, P., Wang, X., Lu, J., Zhang, L., He, Z.: Optimization-based fragmental transmission method for video data in opportunistic networks. Tsinghua Sci. Technol. **4**, 389–399 (2017)
5. Wang, X., Lin, Y., Zhang, S., Cai, Z.: A social activity and physical contact-based routing algorithm in mobile opportunistic networks for emergency response to sudden disasters. Enterp. Inf. Syst. **11**, 597–626 (2017)
6. Zhang, L., Cai, Z., Lu, J., Wang, X.: Mobility-aware routing in delay tolerant networks. Pers. Ubiquitous Comput. **19**, 1111–1123 (2015)

7. Zhang, F., Wang, X., Li, P., Zhang, L.: Energy-aware congestion control scheme in opportunistic networks. IEEJ Trans. Electr. Electron. Eng. **3**, 412–419 (2017)
8. Han, Y., Wu, H., Yang, Z., Li, D.: A new data transmission strategy in mobile D2D networks - deterministic, greedy, or planned opportunistic routing? IEEE Trans. Veh. Technol. **66**, 594–609 (2017)
9. Liu, Y., Wu, H., Xia, Y., Wang, Y., Li, F., Yang, P.: Optimal online data dissemination for resource constrained mobile opportunistic networks. IEEE Trans. Veh. Technol. **66**, 5301–5315 (2017)
10. Wang, S., Wang, X., Cheng, X., Huang, J., Bie, R., Zhao, F.: Fundamental analysis on data dissemination in mobile opportunistic networks with Lévy mobility. IEEE Trans. Veh. Technol. **66**, 4173–4187 (2017)
11. Durisi, G., Koch, T., Popovski, P.: Toward massive, ultrareliable, and low-latency wireless communication with short packets. Proc. IEEE **104**, 1711–1726 (2016)

Detecting Temporal Sentiment-Oriented Difference for Crowdsourced Time-Sync Comments

Ruomiao Li, Yajun Du(✉), and Fei Ren

School of Computer and Software Engineering, Xihua University, Chengdu 610039, China
Ithinkalso@126.com, dyjdoc2003@aliyun.com, 421086313@qq.com

Abstract. Recently, with the rapid development of online shared media, a novel time-sync comment occurs. Time-sync comment is a type of crowdsourced user review embedded in online video website and provides better real-time user interaction than traditional mode. User is probably willing to write comments and then send as time-sync comment when reading interested time-sync comments by others or watching attractive video shots. In this paper, we propose temporal sentiment-oriented difference graph (TSoDG) to detect different areas which exist sentiment-oriented difference over time. We attain the TSoDG by calculating the sentiment-oriented score of time-sync comments, and then use semantic similarity to deal with lagged comments. Finally, we analysis the temporal difference in graph by rule. The result can be regarded as a baseline for further research.

Keywords: Time-sync comment · Sentiment analysis · Difference detection

1 Introduction

More and more online video websites attract a large number of people in recent years. More people spend their time on online video sites for entertainment and write the comments. Comments are not only a user's evaluation of the video itself, but also a way for users to exchange opinions with each other. Now, when watching videos, users would like to have more interaction with other audiences. A novel time-sync comment (TSC; also called bullet-screen comments) occurs and has been applied to many online video websites. In the mechanism, each audience is regarded as a crowdsourced user who can share their feelings or thoughts about video content at any time by sending time-sync comment. These textual contents are organized by video playback time. Users can see the text content sent by other users on the same video playback time. They may be influenced by the existing content, thus bringing richer collision of user ideas. Traditional comment on online video websites is only published after the users watching the whole video. Comparing with traditional way of comment, time-sync comment provides a new way of user interaction. The novel way can lead to large-scale discussions among users.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 318–324, 2020.
https://doi.org/10.1007/978-981-32-9244-4_45

Based on the above phenomenon, we put forward TSoDG, which can express the temporal difference of sentiment-oriented indicated by time-sync comment. Then, we design the method to detect the temporal difference shown in the figure by analyzing the TSoDG. Our main contributions are as follows: (1) we propose a novel TSoDG, which can profile temporal sentiment-oriented difference by time-sync comments. (2) we design a novel method to analysis the temporal difference through TSoDG. (3) In order to better characterize the fluctuation of the period and emphasize difference, we use smoothen method to process the result.

The rest of this article is organized as follows. we first review some previous work about crowdsourced time-sync comments and short text vectorization analysis methods. Then we define TSC-related problem and propose solution to deal with it. In addition, our experiment is executed on a real-world dataset and we analysis the results. Finally, the conclusions of the article are discussed and some possible future work is mentioned.

2 Related Work

At the right moment crowdsourcing provides a suitable way to label data by crowdsource-workers [1]. Time-sync comment is an expression of user comments. In the comment mechanism, users are equivalent to crowdsource-workers. Hence, time-sync comment can be regarded as a kind of crowdsourced data. Shou et al. regarded the work in temporal text summarization as online clustering problem [2]. As we know, a few people researched on time-sync comment since it is a new interactive mode around video sharing. Existing work paid more attention to statistics for comments and study on the relation between comments and videos. Lin et al. proposed a statistical method based on bullet screen text to identify whether a Chinese word comes from a Japanese loanword [3]. Wu et al. studied the relationship between text containing emotional content and video popularity [4]. However, they few worked on content of time-sync comment. In addition, their work was only word-based. It is difficult to guarantee the semantic integrity and unable to accurately obtain the sentiment that users really want to express.

Word embedding techniques were applied in natural language processing, which can more accurately represent the semantic information contained in words. Word or phrases which are represented as real number of vectors in a low dimensional space from vocabulary [5–7]. Our framework uses the representation to analyze the temporal characteristics of the comments. The latest technologies are word2vec [8], GloVe [9] and fastText [10, 11]. In English processing, Bojanowski [12] proposed to use subwords to learn the expression of words. Each word is composed of internal n-gram letters. Compared with the previous method, they considered the micro-deformation characteristics of words. In Chinese processing, Chen et al. [13] proposed a similar method. Grave [14] proposed a method for training such high-quality word representations. Peters et al. [15] introduced a new type of deep context-dependent word representation. The vector of each word is based on the entire text in which it is located.

3 Difference Detection

In this section, according to the temporal features of time-sync comments, we construct the difference graph of temporal sentiment tendency. Meanwhile, we use the semantic features among users' comments to solve the problem of the real-time comment lagging and improve the accuracy of the temporal distribution of sentiment tendency. Then, through analyzing the difference graph, we find the temporal areas showing sentiment differences on the playback time.

3.1 Problem Formulation

As we know, when viewing video, users will comment on the video shots they are interested in. Meanwhile, time-sync comment is a kind of crowdsourced text, users will also see previous comments sent by other users when watching video on the same shot. Thus, the sentiment-oriented varies over time. We define a set of time-sync comments as input, *Comments* = $\{c_1, c_2, \ldots, c_n\}$ with two attributes which are *Timestamp* = $\{t_1, t_2, \ldots, t_n\}$ and *rollDate* = $\{rD_1, rD_2, \ldots, rD_m\}$. In other expression, $c_i(t_i, rD_i)$, i = n. Then we group comments by historical timeline and display sentiment-oriented on the playback time according to timestamps. The illustration is shown in Fig. 1.

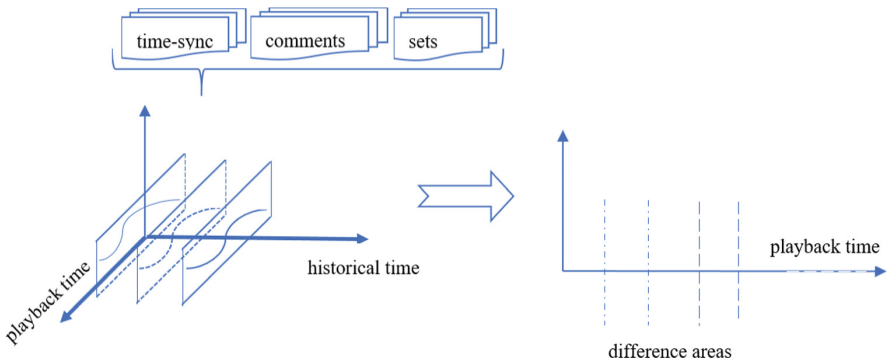


Fig. 1. The illustration of detecting framework

3.2 Temporal Difference Graph Attainment

Statistically, in our data set, the average length of time-sync comments is only 8 Chinese characters. The traditional text analysis method is not suitable for such short text and such short text contains sparse semantics. The state-of-the-art word embedding technique ensure semantic information as much as possible. We use the word embedding method to handle our raw data.

In addition, there has been a challenge. After a user send comment, due to the video continues to play, there is a deviation the point of time that users really want to express and the time when comments appear. In Fig. 2, the user wants to review about the content of scene A. However, scene A only contains few frames. The user writes and

send comment. It displays in the transition between scene A and scene B. We define above the comment as the lagged comment. We use the previous mentioned technologies to solve this problem.

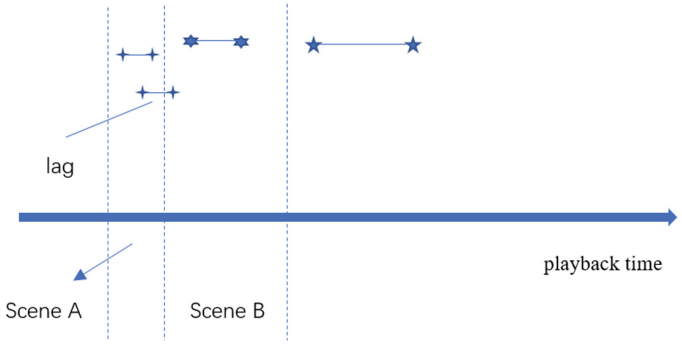


Fig. 2. The illustration of lag-comment

We calculate the mean vector of each word in the sentence, and use this mean vector to represent the vector of the whole sentence. Each vector dimension is 300, so we use cosine similarity to describe the semantic similarity between comments. Use semantic similarity to group lagged comments into the scenarios to which they really belong. *Comment* c_i corresponds to a pair of *Timestamp* t_i and *rollDate* rD_i, where i belongs to the size of *Comment*. Through calculating the sentiment polarity, we attain to the sentiment polarity about c_i which contains a pair of t_i and rD_i. We use t_i and the sentiment polarity of c_i to express the location in graph and describe the variation tendency according to the *rollDate*.

3.3 Difference Judgement

To attain the different areas, we calculate the average affective propensity of the designated playback time t over the global observation period, written as AS_t. We also define different rate (DR_t) to represent the degree of change at playback time t.

$$AS_t = \frac{2}{Z(Z-1)} \sum_{\substack{X,Y \in Z \\ X > Y}} |W_{Xt} - W_{Yt}| \tag{1}$$

Where Z is the whole observation period and W means sentiment polarity weight.

$$DR_t = |St_i - AS_t|, \quad i \in T \tag{2}$$

Where T is time span of the whole observation and i is where playback time t at a specific period.

Besides, to profile the evolution of sentiment over time, we calculate a simple moving average (SMA) across the period to smoothen fluctuation in one playback and emphasize differences across entire times.

4 Experiment

In this section, we conduct experiments on a real-world dataset for detecting the sentiment-oriented difference areas. Our dataset contains 12 episodes, 2,712 video shots, 245,778 comments, 24,186 users. The average user posted 10 comments. We will describe the result of experiments and analysis the evolution of sentiment over time.

In Fig. 3, we present the temporal sentiment-oriented graph of episode11351 (episode id in dataset) which has been smoothed. In order to have a sufficient number of comments to clearly characterize the timing changes, we define the grouping of data according to the timing changes of attention by users after the release of video. For first group the data include the first week and second week. From third week to eighth week, data are classified into a second group. After the ninth week, the remaining data are included in third group.

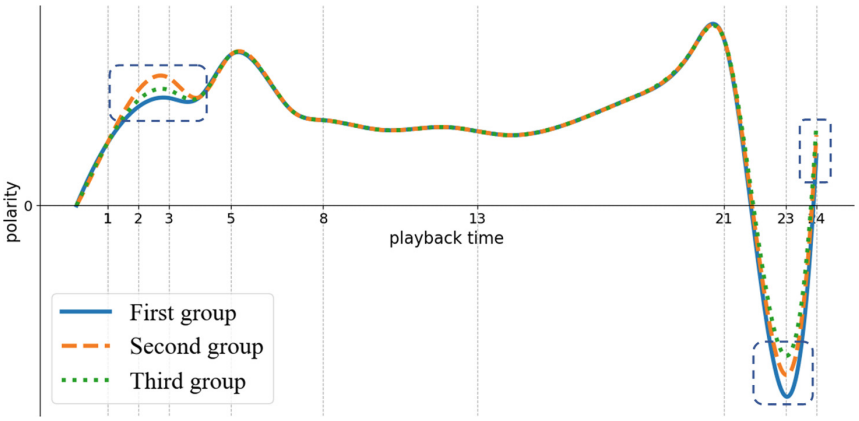


Fig. 3. Difference areas detected

We can clearly see the marked areas which are the ones with temporal differences. From the initial release of video to the Internet, the discussion points of users are all in the first half of video. Especially the introduction part at the beginning, many users watch it several times and have different feelings each time, so the sentiment difference of comments is mainly generated here. With the release of video containing relevant plots, the discussion among users mainly focused on the second half of video. Because the content of video here contains more foreshadowing. The feelings of users who watch it for the first time and other users who watch it again are different. After the ninth week, the story is drawing to a close. Thanks to the beginning and end of the

story, the mystery left by the end of the first episode is gradually clear. Meanwhile, many users have a sentiment resonance at the end. The point can be well reflected in the last difference area.

5 The Conclusion and Future Work

In this paper, we exploit the crowdsourced textual data from time-sync comments to detect the difference areas through the temporal sentiment-oriented graph. In addition, using semantic similarity to process lagged comments. Finally, we find the reason for sentiment evolution by analyzing the result. In the future, we will use the subtitle texts to solve the noise in comments. To improve the detection efficiency and quality, we would like to use the general comments about the story line.

Acknowledgments. This work is one of the projects supported by National Nature Science Foundation (Grant Nos. 61872298, 61532009, 61472329, 61802316 and 61602389). We would like to thank the anonymous reviewers for their helpful comments.

References

1. Archambault, D., Purchase, H.C., Hoßfeld, T.: Evaluation in the crowd: an introduction. In: Archambault, D., Purchase, H., Hoßfeld, T. (eds.) Evaluation in the Crowd. Crowdsourcing and Human-Centered Experiments. LNCS, vol. 10264, pp. 1–5. Springer, Cham (2017)
2. Shou, L., Wang, Z., Chen, K., et al.: Sumblr: continuous summarization of evolving tweet streams. In: International ACM SIGIR Conference on Research & Development in Information Retrieval. ACM (2013)
3. Lin, X., Ito, E., Hirokawa, S., et al.: Chinese tag analysis for foreign movie contents. In: ACIS International Conference on Computer and Information Science. IEEE (2014)
4. Wu, Z., Ito, E.: Correlation analysis between user's emotional comments and popularity measures. In: 3rd International Conference on Advanced Applied Informatics, pp. 280–283 (2014)
5. Turian, J., Ratinov, L., Bengio, Y.: Word representations: a simple and general method for semi-supervised learning. In: 48th Annual Meeting of the Association for Computational Linguistics, pp. 384–394 (2010)
6. Collobert, R., Weston, J., Bottou, L., et al.: Natural language processing (almost) from scratch. J. Mach. Learn. Res. **12**, 2493–2537 (2011)
7. Socher, R., Bauer, J., et al.: Parsing with compositional vector grammars. In: 51st Annual Meeting of the Association for Computational Linguistics, pp. 455–465 (2013)
8. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781 (2013)
9. Pennington, J., Socher, R., Manning, C.D.: GloVe: global Vectors for Word Representation. In: the 2014 Conference on Empirical Methods in Natural Language Processing, pp. 1532–1543 (2014)
10. Joulin, A., Grave, E., Bojanowski, P., Mikolov, T.: Bag of tricks for efficient text classification. arXiv preprint arXiv:1607.01759 (2016)
11. Joulin, A., Grave, E., Bojanowski, P., et al.: Fasttext.zip: compressing text classification models. arXiv preprint arXiv:1612.03651 (2016)

12. Bojanowski, P., Grave, E., Joulin, A., Mikolov, T.: Enriching word vectors with subword information. Trans. Assoc. Comput. Linguist. **5**, 135–146 (2017)
13. Chen, X., Xu, L., Liu, Z., et al.: Joint learning of character and word embeddings. In: The 24th International Joint Conferences on Artificial Intelligence, pp. 1236–1242 (2015)
14. Grave, E., Bojanowski, P., et al.: Learning word vectors for 157 languages. arXiv preprint arXiv:1802.06893 (2018)
15. Peters, M.E., Neumann, M., et al.: Deep contextualized word representations. arXiv preprint arXiv:1802.05365 (2018)

Optimal Control of the Adversarial Information Propagation in Online Social Networks

Xinyan Wang[1,2], Xiaoming Wang[1,2(✉)], Jiehang Xie[1,2], and Pengfei Wan[1,2]

[1] Key Laboratory of Modern Teaching Technology, Ministry of Education, Xi'an 710062, China
[2] School of Computer Science, Shaanxi Normal University, Xi'an 710119, China
wangxm@snnu.edu.cn

Abstract. The popularization of online social networks (OSNs) opens opportunities for information sharing. The existing works mainly investigate the propagation of the single type of information. However, the propagation of adversarial information, that is, the positive and negative information, still remains largely unexplored to date. In this paper, a propagation model is proposed to characterize the propagation dynamics of the adversarial information in OSNs. Afterwards, two different control measures are designed to control the dynamic propagation process and a new control model with control measures is obtained. Then, the minimization of the total system costs is modeled as an optimal control problem. Finally, we obtain an optimal control system using Pontryagin maximum principle. Experiments on the real-world data set show that our proposed control measures is efficient in controlling the propagation of the adversarial information.

Keywords: Online social networks · Adversarial information · Optimal control

1 Introduction

With the popularity of the social network platforms, information sharing has become prevalent for individuals [1]. On one hand, the propagation of the positive information brings great convenience in many aspects, such as health care and education. Unfortunately, the propagation of the negative information, such as rumor and gossip has adverse impacts on individual's decisions, which can cause social panic, anxiety and crushes. Therefore, a reasonable propagation model for characterizing the propagation dynamics of the adversarial information and efficient control measures must be developed to decrease the influence of the negative information propagation. Answering these questions can provide the guidance for controlling the propagation of the adversarial information.

The existing works involved the Linear Threshold (LT) model [2] and the Independent Cascade (IC) [3] model mainly utilize these models to study the propagation process of the information based on individuals. Besides, the Susceptible-Infected-Recovered

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 325–331, 2020.
https://doi.org/10.1007/978-981-32-9244-4_46

(SIR) model [4] in biology also is utilized to describe the propagation dynamics. For example, [5] proposed a delayed SLBS computer virus model and investigated the existence of the optimal control strategy to control the propagation of the computer virus. [6] established a coupled multiplex network to model the propagation of epidemic dynamics using Markov chain approach. In addition, [7] developed a rumor spreading model considering psychological factors time delay. [8] established a model to study the simultaneous propagation of the negative and positive information.

However, aforementioned studies mainly focus on the propagation of the single type of information. Meanwhile, individual's decision-making behaviors are influenced by psychology, emotion etc. Consequently, how to model the propagation of the adversarial information and control the dynamic propagation process is a challenging work.

In this paper, we aim to establish a propagation model of the positive and negative information to characterize the dynamic propagation process of the adversarial information. Our main contribution is three-fold: (1) We divide the states presented by different individuals into four types and then propose a propagation model, called *SNPR* to characterize the dynamic propagation process of the adversarial information. Then, we design two different control measures to decrease the proportion of individuals affected by the negative information. (2) We model the minimization of the total costs as an optimal control problem. Afterwards, we derive the dynamic change of two control measures over time and obtain an optimal control system. (3) We conduct experiments on real-world data set to validate the efficiency of the control measures. Experimental results demonstrate that our proposed control measures can reduce the proportion of individuals propagated the negative information apparently.

The remainder of this work is organized as follows: Sect. 2 is our system model. In Sect. 3, the optimal control measures are presented. We conduct experiments in Sect. 4. The paper is concluded in Sect. 5.

2 System Model

In this section, for the sake of characterizing the dynamic propagation process of the adversarial information, we propose a propagation model, called *SNPR* of the positive and negative information.

In view of the complicated interaction relationships among individuals, when the adversarial information is propagated in the network, the individuals received information are usually divided into four states: Susceptible state (S) represents that the individual holds neither the positive information, nor the negative information yet; Negative information-propagated state (N) represents that the individual receives the negative information and then propagate it; Positive information-propagated state (P) represents that the individual receives the positive information and then propagate it; Mixed information-chaotic state (R) represents that the individual holds both the positive information and the negative information simultaneously.

Supposed that any individual is in one of four stats at any time and all individuals in four different states may leave the network with probability ω. In what follows, we can define the following transition relationships of different individuals among four states:

(1) If an individual in state S receives the positive information, he transfers to the state P with probability β; or if an individual in state S receives the negative information, he transfers to the state N with probability α. (2) If an individual in state N receives the positive information, this individual transfers to state R with probability φ. (3) If an individual in state P receives the negative information, this individual transfers to state R with probability δ. (4) If an individual in state R receives the negative information, he transfers to state N with probability λ; or if an individual in state R receives the positive information, he transfers to state P with probability μ.

Here, we let $S(t), N(t), P(t), R(t)$ denote the number of individuals in four state, respectively. Additionally, we let Λ denotes the total number of individuals. Similarly, $s(t), n(t), p(t), r(t)$ denote the proportion of individuals in four states, respectively, where $s(t) + n(t) + p(t) + r(t) = 1$. At the same time, the total number of individuals is unchanging, that is, $S(t) + N(t) + P(t) + R(t) = \Lambda$.

Based on above transition relationships, the *SNPR* model is formulated as follows:

$$ds(t)/dt = \omega - \alpha s(t)n(t) - \beta s(t)p(t) - \omega s(t) \tag{1}$$

$$dn(t)/dt = \alpha s(t)n(t) + \lambda r(t)n(t) - \varphi n(t)r(t) - \omega n(t) \tag{2}$$

$$dp(t)/dt = \beta s(t)p(t) + \mu r(t)p(t) - \delta p(t)r(t) - \omega p(t) \tag{3}$$

$$dr(t)/dt = \varphi n(t)r(t) + \delta p(t)r(t) - \lambda r(t)n(t) - \mu r(t)p(t) - \omega r(t) \tag{4}$$

where $s(0) \geq 0,\ n(0) \geq 0,\ p(0) \geq 0, r(0) \geq 0$. As presented above, the propagation process of the adversarial information is investigated. Besides, we propose two control measures to inhibit the propagation of the adversarial information.

3 Optimal Control Measures

In this section, we design two different control measures. In other words, $v_1(t)$ and $v_2(t)$ represent the control measures. Now, we consider an optimal control problem, which is the minimization of the objective function:

$$J = \int_0^T \left[an(t) + br(t) + mv_1^2(t) + nv_2^2(t)\right] dt \tag{5}$$

subject to:

$$ds(t)/dt = \omega - \alpha s(t)n(t) - \beta s(t)p(t) - \omega s(t) \tag{6}$$

$$dn(t)/dt = \alpha s(t)n(t) + \lambda r(t)n(t) - \varphi n(t)r(t) - \omega n(t) - \varepsilon_1 n(t)v_1(t) \tag{7}$$

$$dp(t)/dt = \beta s(t)p(t) + \mu r(t)p(t) - \delta p(t)r(t) - \omega p(t) + \varepsilon_1 n(t)v_1(t) + \varepsilon_2 r(t)v_2(t) \tag{8}$$

$$dr(t)/dt = \varphi n(t)r(t) + \delta p(t)r(t) - \lambda r(t)n(t) - \mu r(t)p(t) - \omega r(t) - \varepsilon_2 r(t)v_2(t) \quad (9)$$

where T is the control time; a and b denote the losses generated by individuals in state N and R, respectively. m and n denote the weight of the two control measures.

In order to seek the optimal control measures, we should first define the Lagrangian function and Hamiltonian function of this optimal control problem, in which the Lagrangian function is given by:

$$L = an(t) + br(t) + mv_1^2(t) + nv_2^2(t) \quad (10)$$

Afterwards, we also can define the corresponding Hamiltonian function as:

$$H = L + \varsigma_s(t)ds(t)/dt + \varsigma_n(t)dn(t)/dt + \varsigma_p(t)dp(t)/dt + \varsigma_r(t)dr(t)/dt \quad (11)$$

Combining the (6)–(11), we can obtain equation given by (12):

$$\begin{aligned} H = {} & an(t) + br(t) + mv_1^2(t) + nv_2^2(t) + \varsigma_s(t)[\omega - \alpha s(t)n(t) - \beta s(t)p(t) - \omega s(t)] \\ & + \varsigma_n(t)[\alpha s(t)n(t) + \lambda r(t)n(t) - \varphi n(t)r(t) - \omega n(t) - \varepsilon_1 n(t)v_1(t)] \\ & + \varsigma_p(t)[\beta s(t)p(t) + \mu r(t)p(t) - \delta p(t)r(t) - \omega p(t) + \varepsilon_1 n(t)v_1(t) + \varepsilon_2 r(t)v_2(t)] \\ & + \varsigma_r(t)[\varphi n(t)r(t) + \delta p(t)r(t) - \lambda r(t)n(t) - \mu r(t)p(t) - \omega r(t) - \varepsilon_2 r(t)v_2(t)] \end{aligned} \quad (12)$$

where $\varsigma_s(t), \varsigma_n(t), \varsigma_p(t), \varsigma_r(t)$ is the adjoint functions. Besides, the optimal control measures are given by Theorem 1 and $0 \le v_1(t) \le 1,\ 0 \le v_2(t) \le 1$.

Theorem 1. Given an optimal control $v^*(t) \in \{v_1^*(t), v_2^*(t)\}$ and the corresponding states $s^*(t),\ n^*(t),\ p^*(t),\ r^*(t)$ of the states, there must exist adjoint functions $\varsigma(t) \in \{\varsigma_s(t), \varsigma_n(t), \varsigma_p(t), \varsigma_r(t)\}$ that satisfying:

$$d\varsigma_s^*(t)/dt = \varsigma_s^*(t)[\alpha n(t) + \beta p(t) + \omega] - \varsigma_n^*(t)\alpha n(t) - \varsigma_p^*(t)\beta p(t) \quad (13)$$

$$\begin{aligned} d\varsigma_n^*(t)/dt = {} & \varsigma_s^*(t)\alpha s(t) - \varsigma_n^*(t)[\alpha s(t) + (\lambda - \varphi)r(t) - \omega - \varepsilon_1 v_1(t)] \\ & + \varsigma_p^*(t)[\mu r(t) + \varepsilon_1 v_1(t)] + \varsigma_r^*(t)(\varphi - \lambda)r(t) \end{aligned} \quad (14)$$

$$d\varsigma_p^*(t)\Big/dt = \varsigma_s^*(t)\beta s(t) - \varsigma_p^*(t)[\beta s(t) - \delta r(t) - \omega] + \varsigma_r^*(t)(\mu - \delta)r(t) \quad (15)$$

$$\begin{aligned} d\varsigma_r^*(t)/dt = {} & \varsigma_n^*(t)(\varphi - \lambda)n(t) + \varsigma_p^*(t)[\delta p(t) - \mu n(t) - \varepsilon_2 v_2(t)] \\ & + \varsigma_r^*(t)[(\lambda - \varphi)n(t) + (\mu - \delta)p(t) + \omega + \varepsilon_2 v_2(t)] \end{aligned} \quad (16)$$

with transversality conditions: $\varsigma_s(T) = \varsigma_n(T) = \varsigma_p(T) = \varsigma_r(T) = 0$.

Besides, the optimal control measures $v_1^*(t)$ and $v_2^*(t)$ are given by:

$$v_1^*(t) = \max\{\min\{x, v_{1_{\max}}\}, 0\} \tag{17}$$

$$v_2^*(t) = \max\{\min\{y, v_{2_{\max}}\}, 0\} \tag{18}$$

Proof. There must exist a nontrivial vector function $\varsigma^*(t)$ satisfying the following conditions: $d\varsigma(t)/dt = -\partial H/\partial q(t)$ and $\partial H/\partial v(t) = 0$.

Then, we differentiate the Hamiltonian function with respect to the state variables $q(t) \in \{s(t),\ n(t),\ p(t),\ r(t)\}$ to obtain the adjoint functions using Pontryagin maximum principle [9]. Thus, we can obtain the following equations:

$$\partial H/\partial v_1(t) = 2mv_1^*(t) - \varsigma_n(t)\varepsilon_1 n(t) + \varsigma_p(t)\varepsilon_1 n(t) = 0 \tag{19}$$

$$\partial H/\partial v_2(t) = 2nv_2^*(t) + \varsigma_p(t)\varepsilon_2 r(t) - \varsigma_r(t)\varepsilon_2 r(t) = 0 \tag{20}$$

Considering the boundary conditions of the optimal control measures, we have:

$$v_1^*(t) = \begin{cases} 0, & if\ x < 0 \\ v_{1_{\max}}, & if\ x > v_{1_{\max}} \\ x, & otherwise \end{cases} \tag{21}$$

$$v_2^*(t) = \begin{cases} 0, & if\ y < 0 \\ v_{2_{\max}}, & if\ y > v_{2_{\max}} \\ y, & otherwise \end{cases} \tag{22}$$

where $x = \varepsilon_1 n(t)\left[\varsigma_n(t) - \varsigma_p(t)\right]/2m$, $y = \varepsilon_2 r(t)\left[\varsigma_r(t) - \varsigma_p(t)\right]/2n$; In addition, $v_{1_{\max}}, v_{2_{\max}}$ denote the upper bound of $v_1(t)$ and $v_2(t)$. We rewrite the expression (21)–(22) as in (17)–(18). Thus, we have finished the proof of Theorem 1. After substituting (17) and (18) into (6)–(9) and (13)–(16), we can obtain an optimal control system.

4 Experiments

In this section, we evaluate the efficiency of control measures in Matlab, using a Intel Core(TM) i5-7200U CPU. Besides, we utilize Slashdot dataset, which is a technology-related news website allowing users to tag each other as "friends" or "foes", a signed link (a friend relationship) means that a user likes another user's comments, while a foe relationship means that a user finds another user's comments uninteresting. The Slashdot network is composed of 82,144 users and 549,202 edges [10]. The basic parameters are as follows: $\omega = 0.01, a = 0.45, b = 0.2, m = 0.6, n = 0.48$.

The experimental results in Figs. 1, 2, 3 and 4 reveal the dynamic propagation process of the positive and negative information. As shown in Fig. 1, the two control measures are not implemented at all, the maximum proportion of individuals in states N and P is 30% and 52%, respectively.

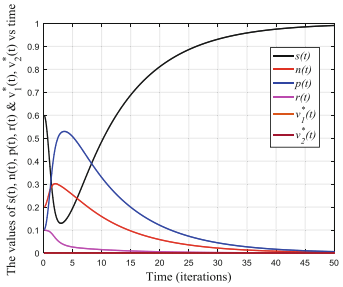


Fig. 1. There is no control measure

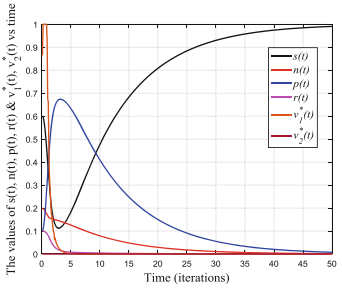


Fig. 2. There is only $v_1(t)$ measure

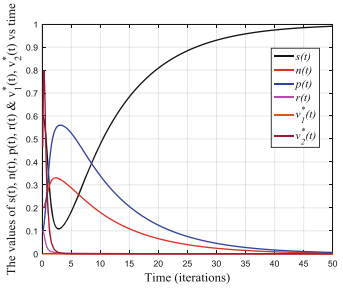


Fig. 3. There is only $v_2(t)$ measure

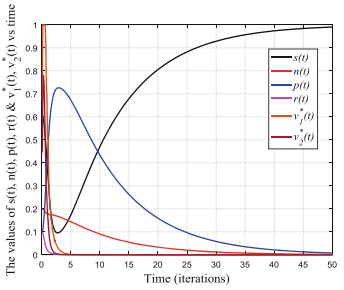


Fig. 4. There is $v_1(t)$ & $v_2(t)$ measure

Additionally, when one control measure is implemented only in Figs. 2 and 3, the proportion of individuals propagated the positive information is 68% and 56%, respectively. More importantly, when two control measures are implemented simultaneously in Fig. 4, the proportion of individuals propagated the positive information is 72%, which is improved by 40% approximately.

Therefore, the performance of our proposed control measures is superior when two control measures are implemented simultaneously.

5 Conclusion

This paper proposes a propagation model of the adversarial information in OSNs to describe the dynamic propagation process of the positive and negative information. Then, two control measures are designed and an optimal control system is obtained. The experimental results show that out proposed control measures can decrease the proportion of individuals affected by the negative information propagation.

Acknowledgments. This work is supported by the National Natural Science Foundation of China (Grant No. 61872228).

References

1. Ji, F., Tay, W.P.: An algorithmic framework for estimating rumor sources with different start times. IEEE Trans. Signal Process. **65**(10), 2517–2530 (2017)
2. Bozorgi, A., Samet, S., Kwisthout, J., et al.: Community-based influence maximization in social networks under a competitive linear threshold model. Knowl.-Based Syst. **134**, 149 (2017)
3. Wang, Q., Jin, Y., Lin, Z., et al.: Influence maximization in social networks under an independent cascade-based model. Physica A **444**, 20–34 (2016)
4. Zhu, K., Ying, L.: Information source detection in the SIR model: a sample path based approach. IEEE/ACM Trans. Netw. **24**(1), 408–421 (2012)
5. Chen, L., Hattaf, K., Sun, J.: Optimal control of a delayed SLBS computer virus model. Physica A **427**, 244–250 (2015)
6. Zheng, C., Xia, C., Guo, Q., et al.: Interplay between SIR-based disease spreading and awareness diffusion on multiplex networks. J. Parallel Distrib. Comput. **115**, 20–28 (2018)
7. Huo, L., Ma, C.: Optimal control of rumor spreading model with consideration of psychological factors and time delay. Discrete Dynam. Nature Soc. **2018**, 1–12 (2018)
8. Wen, S., Haghighi, M.S., Chen, C., et al.: A sword with two edges: propagation studies on both positive and negative information in online social networks. IEEE Trans. Comput. **64** (3), 640–653 (2015)
9. Ali, H.M., Pereira, F.L., Gama, S.M.A.: A new approach to the Pontryagin maximum principle for nonlinear fractional optimal control problems. Math. Methods Appl. Sci. **39** (13), 3640–3649 (2016)
10. Leskovec, J., Huttenlocher, D., Kleinberg, J.: Signed networks in social media. In: 28th ACM Conference on Human Factors in Computing Systems (CHI) (2010)

A Hierarchical Attention Headline Generation Model with a Filter

Jiehang Xie[1,2], Xiaoming Wang[1,2(✉)], Xinyan Wang[1,2], and Guangyao Pang[1,2]

[1] Key Laboratory of Modern Teaching Technology, Ministry of Education, Xi'an 710062, China
wangxm@snnu.edu.cn
[2] School of Computer Science, Shaanxi Normal University, Xi'an 710119, China

Abstract. The recent researches demonstrate the methods of generating text headline can tackle the problem of information overload. However, the popular neural network based framework, namely, encoder-decoder framework is difficult to handle the headline generation task of long source text. In this paper, we establish a hierarchical attention headline generation model with a filter to solve this problem. This model first relies on a filter to extract crucial contents of the source text, and then the hierarchical attention mechanism accurately identifies important words. Finally, our model generates a high-quality headline. Experimental results show that the ROUGE scores of our model are higher than those of classical models. Furthermore, our model achieves better performance than classical model in dealing with long text.

Keywords: Headline generation · Attention mechanism · Neural network

1 Introduction

The purpose of the headline generation is to generate a text headline with only one sentence for a source text, which preserves the most important information of the source text. Compressing a large amount of text information into headlines can effectively solve the problem of information overload [1].

Traditionally, the text headline generation methods can be categorized into two types: extractive and abstractive. Among them, extractive method obtains a headline by extracting the original sentence of the source text. For example, [2] integrated fuzzy logic, bushy path and word-net synonyms to identify important sentences in the source text. Besides, [3] utilized the multidimensional knowledge representation framework to extractive headline. By comparison, abstractive method generates a headline using words or sentences that have not appeared in the source text. In fact, some recent studies demonstrate that the text headline generated by abstractive method is more similar to that of manual generation [4]. Therefore, the abstractive method has attracted widely concern of researchers. [5] established a encoder-decoder framework based on neural network to generate headline, which has made impressive performance. [6] extended the encoder-decoder framework, selecting important sentences and words

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 332–338, 2020.
https://doi.org/10.1007/978-981-32-9244-4_47

from the source text before decoding. [7] used a dual encoder model to encode the source text coarsely and finely, respectively.

Unfortunately, there are several issues remain to be solved in the aforementioned models. For example, the popular encoder-decoder framework is more effective only when handling short texts with less characteristics, resulting in poor performance in generating headline for long source texts.

To address this issue, we propose a hierarchical attention headline generation model with a filter, called *HAF*. Specifically, four classical extractive methods with good performance are used to construct the filter in this model to obtain the important parts of the source text, called coarse summaries. Following, these coarse summaries are fed into a encoder-decoder model based on hierarchical attention framework to generate a accurate text headline. Consequently, this gives the encoder-decoder attention framework the ability to generate high-quality headline for long source text. Finally, for the sake of evaluating the performance of our model, we conduct experiments on the Gigaword and DUC-2004 data sets, the experimental results demonstrate that the performance of our model is significantly superior to that of the classical headline generation models.

The remainder of this paper is organized as follows: Sect. 2 presents our problem formulation. Section 3 describes our model. The experimental evaluation is conducted in Sect. 4. Finally, Sect. 5 concludes this paper.

2 Problem Formulation

In this section, we first illustrate the flaws of the common encoder-decoder model using an example. Then, we present the problem formulation of headline generation. Obviously, as shown in Table 1, we can observe that the model is able to capture the main information of the short source text and generate a headline that conforms to the facts. By contrast, when the source text is long, the model loses the main information of the source text and generates a ridiculous headline.

Definition 1: In order to generate a good headline, The headline generation model find y_1 that maximize the conditional probability p under the condition of X on the premise of the constraints in Eq. (2), which can be shown as follows:

$$Y = \arg\max_{y_t} \prod_{t=1}^{n} p(y_t \mid \{y_1, \ldots, y_{t-1}\}, X) \tag{1}$$

$$s.t. \begin{cases} n < m \\ x_i, y_j \in V \\ 1 \le i \le m, 1 \le j \le n \end{cases} \tag{2}$$

where X is the input source text content $X = \{x_1, x_2, \ldots, x_m\}$, m is the length of the source text, x_i denotes the i-th word in the source text, V is the word vocabulary, and the generated headline is represented by $Y = \{y_1, y_2, \ldots, y_n\}$, in which the j-th word is denoted as y_j, n is the word number of the headline.

3 Model

In what follows, we describe our proposed HAF model, which consists of a filter and an encoder-decoder framework with hierarchical attention mechanism.

Table 1. An example of generating headlines for source text of different lengths

Short text	Four children were killed and another three wounded thursday when an old mortar fuse exploded as they played with it in Afghanistan's capital, police said
Standard headline	Children killed by old explosives in Afghanistan
Generated headline	Four children killed in northern Afghanistan
Long source text	It is a cast worthy of a political thriller: a former convict whose claims about a former president and first lady spurred a criminal investigation; a prominent senator's brother-in-law, who worked undercover for federal agents looking into the case; sworn enemies of the former first family trying to dig up fresh dirt, and some hollywood stars
Standard headline	Trial over hillary clinton's hollywood event has a strong cast
Generated headline	Former white house shows banda's first lady

3.1 The Filter

For the sake of solving the problem of information loss in common encoder-decoder model, we construct a filter based on four different mature extractive methods. The filter extracts important parts of the source text and obtain the coarse summary $\{S_1, S_2, S_3, S_4\}$, the maximum length of these coarse summaries is no more than 45 words. The four extractive methods are as follows:

(1) *TextRank* is a graph-based sorting algorithm for text. The basic idea comes from Google's PageRank algorithm [8].
(2) *LexRank* is also a graph-based method. The method uses the similarity of sentences to calculate the weight of sentences [9].
(3) *SumBasic* is a word-frequency based summarization system [10].
(4) *KL-sum* takes advantage of topic vocabulary distribution to identify important contents in source text [9].

3.2 Encoder-Decoder

The encoder is used to map input source text into a vector representation, that is, the context vector. In our model, the encoder accepts four coarse summaries and generates four context vectors. We adopt a bidirectional Gated Recurrent Unit(GRU) [11]

encoder to capture dependencies of source text and increase the integrity of the information contained in the context vector. A bidirectional GRU consists of forward and backward GRU.

At time step t, the hidden state h_t^e is concatenated by the forward hidden state $\overrightarrow{h}_t^e$ with the backward hidden state $\overleftarrow{h}_t^e$, which are calculated as follows:

$$\overrightarrow{h}_t^e = GRU\left(x_t, \overrightarrow{h}_{t-1}^e\right) \tag{3}$$

$$\overleftarrow{h}_t^e = GRU\left(x_t, \overleftarrow{h}_{t-1}^e\right) \tag{4}$$

$$h_t^e = \left[\overrightarrow{h}_t^e : \overleftarrow{h}_t^e\right] \tag{5}$$

Afterwards, the context vector c is obtained as the average of all the encoder hidden states: $c = \left(\sum_{i=1}^{u} h_i^e\right)/u$, where u represents the length of coarse summaries. Ultimately, we integrate four context vectors $\{c_1, c_2, c_3, c_4\}$ to acquire the composite context vector: $c^s = \left(\sum_{o=1}^{p} c_o\right)/p$, where p is the number of coarse summaries.

The attention decoder accepts the composite context vector as the initial state, and then determines which coarse summaries and word should be concerned. Finally, the attention decoder generates a headline word. To be specific, the hidden state of the decoder is determined by the output and hidden state of the previous time step:

$$h_t^d = GRU\left(y_{t-1}, h_{t-1}^d\right) \tag{6}$$

Additionally, the attention weight s^l of coarse summaries is calculated as follows:

$$s^l = \exp(c_l \cdot h_t^d) / \sum_{o=1}^{p} \exp(c_o \cdot h_t^d) \tag{7}$$

where h_t^d represents hidden state of the decoder at the time step t.

The attention weight ω_{oi}^t of the word in the coarse summaries is computed as:

$$\omega_{oi}^t = \exp(h_{oi}^e \cdot h_t^d) / \sum_{j=1}^{m} \exp(h_{oj}^e \cdot h_t^d) \tag{8}$$

where h_{oi}^e is the i-th hidden state in the o-th coarse summaries.

Finally, we predict the headline word based on the context weight a^t and the hidden state of the current time step:

$$a^t = \sum_{o=1}^{p} \sum_{i}^{m} s^l \omega_{oi}^t h_{oi}^e \tag{9}$$

$$p(y_t \mid y_1, \ldots, y_{t-1}) = \varsigma\left(W_{h^d} h_t^d + W_a a^t\right) \tag{10}$$

where $\varsigma(\cdot)$ is the normalized exponential function. In the test process, we utilize beam search algorithm with beam size of 5 for decoding.

4 Experimental Evaluation

We train our model with Gigawords corpus, which is an English sentence summary dataset containing texts and headlines from many information publishing platforms such as The Xinhua News Agency English Service and Agence France Press English Service, DUC-2004 is regarded as test set, which has 500 news from New York Times and Associated Press Wire. Besides, We verify the performance of our model, by comparing with the following benchmark models. *PRE* is an extractive method that outputs the first 75 characters of the source text as a headline. *ABS* uses CNN encoder and NNLM decoder to complete the headline generation task [5]. *S2S-A* is a common encoder-decoder model based on standard attention mechanism. Sum-Hi [6] is a neural network model based on coarse-to-fine attention mechanism. In addition, we take the widely used ROUGE [12] in headline quality evaluation as evaluation metrics. ROUGE is an automatic summary evaluation method for evaluating recall rate of n-gram. It evaluates abstracts based on co-occurrence information of n-gram in abstracts.

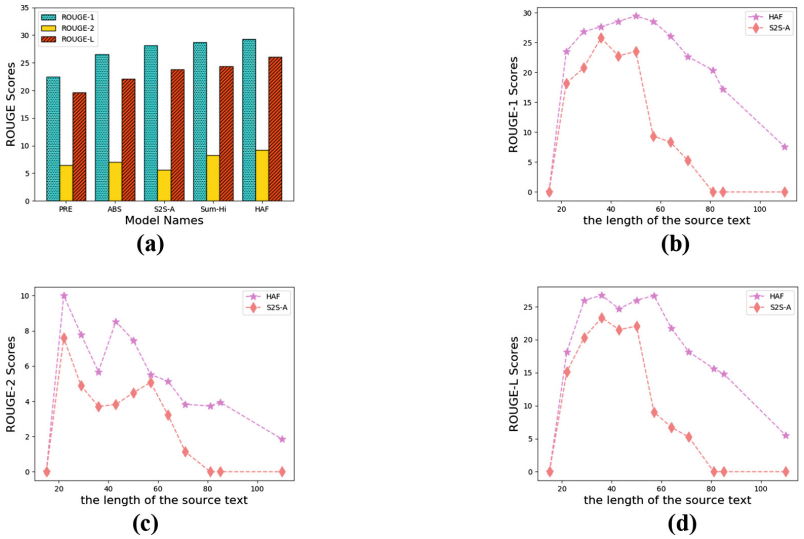


Fig. 1. ROUGE scores of our model and baseline model.

As shown in Fig. 1(a), the ROUGE-1, ROUGE-2 and ROUGE-L of our model are 2.72, 1.51 and 4.01 higher than those of the ABS model, respectively, which means that the quality of the headline generated with our model is superior to that of the baseline model.

Figure 1(b–d) reveal that the ROUGE scores obtained by our model is apparently higher than that of the SAS model on source texts of different lengths, which demonstrate that the HAF model can capture the important information of long text more accurately. Therefore, the performance of our model is superior to that of the SAS model.

5 Conclusion

In this paper, we propose a hierarchical attention headline generation model with a filter to extract important parts of the source text to obtain coarse summaries. Then, the coarse summary is input into a encoder-decoder framework with the hierarchical attention mechanism to generate a headline. Experiments on DUC-2004 dataset reveal that our model can generate a high-quality headline and solve the problem of information loss. In our future work, we plan to focus our efforts on improving our model performance.

Acknowledgments. This work was supported by the National Natural Science Foundation of China [Grant Nos. 61872228].

References

1. Yao, J., Wan, X., Xiao, J.: Recent advances in document summarization. Knowl. Inf. Syst. **53**(2), 297–336 (2017)
2. Yadav, J., Meena, Y.K.: Use of fuzzy logic and wordnet for improving performance of extractive automatic text summarization. In: Advances in Computing, Communications and Informatics, pp. 2071–2077. IEEE (2016)
3. Zenkert, J., Klahold, A., Fathi, M.: Towards extractive text summarization using multidimensional knowledge representation. In: IEEE International Conference on Electro/Information Technology, pp. 0826–0831. IEEE (2018)
4. Bing, L., Li, P., Liao, Y., Lam, W., Guo, W., Passonneau, R.J.: Abstractive multi-document summarization via phrase selection and merging. In: The 53rd Annual Meeting of the Association for Computational Linguistics, pp. 1587–1597 (2015)
5. Rush, A.M., Chopra, S., Weston, J.: A neural attention model for abstractive sentence summarization. In: The Conference on Empirical Methods in Natural Language Processing, pp. 379–389 (2015)
6. Tan, J., Wan, X., Xiao, J.: From neural sentence summarization to headline generation: a coarse-to-fine approach. In: The 26th International Joint Conference on Artificial Intelligence, pp. 4109–4115. AAAI Press (2017)
7. Yao, K., Zhang, L., Du, D., Luo, T., Tao, L., Wu, Y.: Dual encoding for abstractive text summarization. IEEE Trans. Cybern., 1–12 (2018)
8. Mihalcea, R., Tarau, P.: Textrank: bringing order into text. In: The Conference on Empirical Methods in Natural Language Processing, pp. 404–411 (2004)

9. Haghighi, A., Vanderwende, L.: Exploring content models for multi-document summarization. In: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, pp. 362–370 (2009)
10. Vanderwende, L., Suzuki, H., Brockett, C., et al.: Beyond SumBasic: task-focused summarization with sentence simplification and lexical expansion. Inf. Process. Manage. **43** (6), 1606–1618 (2007)
11. Cho, K., Van Merriënboer, B., Gulcehre, C., et al.: Learning phrase representations using RNN encoder-decoder for statistical machine translation. In: The Conference on Empirical Methods in Natural Language Processing, pp. 1724–1734 (2014)
12. Lin, C.Y.: Looking for a few good metrics: rouge and its evaluation. In: Proceedings of the Evaluation of Information Access Technologies (2004)

An ICP-Based Point Clouds Registration Method for Indoor Environment Modeling

Su Sun[1], Wei Song[1(✉)], Yifei Tian[2], and Simon Fong[2]

[1] North China University of Technology, No. 5 Jinyuanzhuang Road, Shijingshan District, Beijing 100-144, China
sunsu71@foxmail.com, sw@ncut.edu.cn
[2] Department of Computer and Information Science, University of Macau, Macau, China
tianyifei0000@sina.com, ccfong@umac.mo

Abstract. LiDAR has been widely used in 3D reconstruction due to its high resolution, wide range and tolerance towards light and weather. To realize accurate and complete environment perception and reconstruction, LiDAR point cloud registration plays a crucial role. This paper utilized an Iterative Closest Point (ICP) algorithm to register the sparse point cloud sensed by LiDAR into a whole indoor environment model. Instead of using a standard ICP algorithm, a point-to-plane ICP is adopted with point cloud selection, point pair matching and rejection. The transformation value between two point cloud data is iteratively calculated and optimized until the defined error metric reaches convergence.

Keywords: LiDAR · 3D point cloud · Environment modeling · ICP

1 Introduction

Currently, 3D reconstruction technology is a hot topic in environment perception and path planning of mobile robots, and has been widely studied and applied in the fields of science and engineering. In the environmental reconstruction of mobile robot, a variety of sensors, such as radar and video cameras, were utilized to sense the environmental information to achieve 3D terrain reconstruction. In image-based reconstruction, due to the resolution limit and the vulnerability to weather and light of image sensors, it was hard to obtain sufficient 3D environmental information [1]. Moreover, in the absence of obvious characteristics of the environment, single image sensors were not conducive to achieving accurate terrain reconstruction. In recent years, due to the development of LiDAR, combined with graphic technology, terrain reconstruction from 2D to 3D has been further improved. LiDAR can obtain accurate 3D information of the environment based on laser ranging, which is not affected by external factors such as light and weather. Therefore, LiDAR is widely used in 3D terrain reconstruction.

Using LiDAR to perceive the environment, point cloud data collected at a single time is often insufficient due to the limitation of radar detection range and the occlusion of objects in the environment. Therefore, it is necessary to collect multiple point clouds from different locations and register each point cloud into one global terrain, so as to

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 339–344, 2020.
https://doi.org/10.1007/978-981-32-9244-4_48

achieve more accurate and complete environment reconstruction. Different methods have been addressed to implement registration, based on features, objects, and descriptors. Jaw [2] extracted point, line, and plane features, from the scene, as a combined feature to estimate the transformation parameters between two point clouds, on the basis of which point cloud registration was performed. Boulc et al. [3] firstly divided point clouds into several separate objects and calculated correspondence between objects. After the correspondence was determined, key points were matched based on object features, instead of only geometric features, for subsequent point cloud registration.

In point clouds registration, ICP algorithm is the most famous method for calculating optimal transformation between two point clouds, by iteratively minimizing defined error metrics. The most commonly used ICP algorithm is the point-to-point ICP proposed by Besl and McKay [5], which calculated the transformation matrix that optimizes point-to-point metric based on Euclidean distance between corresponding points. However, in LiDAR point cloud registration, traditional point-to-point ICP is hard to obtain ideal registration results, because collected point clouds are sparse and inhomogeneous, and the physical characteristics of measured points in two point clouds are usually inconsistent. Chen et al. [6] improved the point-to-point ICP by substituting a point-to-plane metric for Euclidean distance. The distance between a point and its corresponding plane was optimized along the surface normal, which improved the registration significantly.

In this paper, we utilized the point-to-plane ICP method to calculate and optimize the transformation value between input data, thus to perform the accurate registration of indoor environmental point clouds. The remainder of this paper is organized as follows. Section 2 introduces the framework of the proposed registration method. Section 3 displays the result of the experiment. Section 4 makes a conclusion and presents the future work.

2 The ICP-Based Registration Method of Point Clouds

The ICP algorithm is the most prevalent method for executing accurate registration of point clouds. The purpose of ICP algorithm is to find the correspondence between point clouds in different frames by iteratively calculating the optimal rigid body transformation, including translation and rotation value, until a defined error metric reaches convergence. The process of an ICP algorithm includes applying filters to select preferential points from input point clouds, matching point pairs from source point cloud and target point cloud, discarding invalid point pairs, and minimizing an error metric to calculate the optimal correspondence between two input point clouds. Figure 1 shows the flowchart of the ICP algorithm.

Proper selection improves the efficiency of ICP in processing mass point cloud data, especially dense point clouds, but as the point cloud collected by LiDAR is sparse, selection will further reduce information and features of obtained point cloud data, causing poor registration results. Therefore, we only apply a height threshold to filter ground points, considering the interference caused by ground points when matching point pairs and calculating point features, and the rest points are all utilized

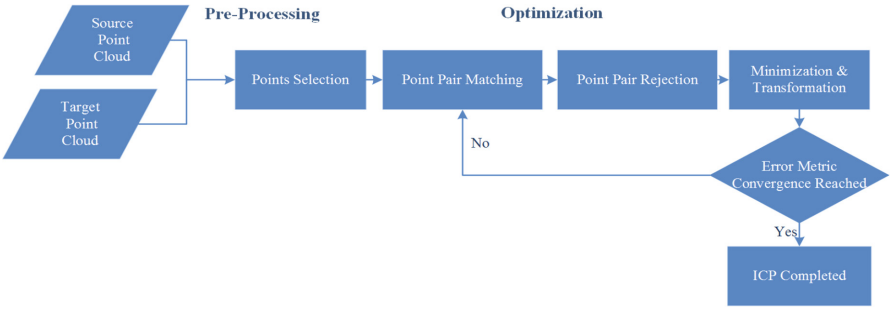
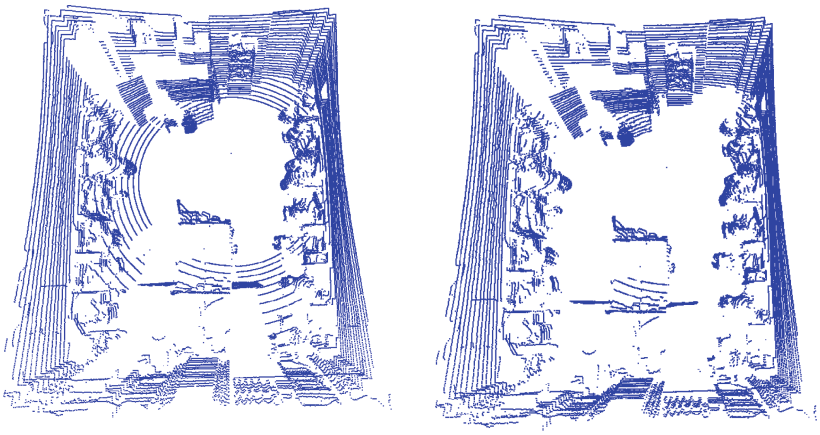


Fig. 1. The Flowchart of ICP algorithm

without selecting and sampling to avoid information loss. Since the point cloud data tested in this paper was collected indoors, with almost no floor tilt, the height threshold is simply defined based on the distance in vertical direction. Figure 2 shows the result after filtering ground points.



(a) (b)

Fig. 2. Example of the Point Filtering: (a) raw point cloud; (b) the same point cloud data with ground points removed

Point matching in ICP finds point pairs between source points and target points, on the basis of which transformation information is calculated and optimized. The correspondence pairs are obtained from searching the nearest point in target point cloud for each point in source point cloud. The nearest point finding strategy in this paper is determined by the Euclidean distance between points. Meanwhile, wrong point pairs can significantly influence the result of calculating correspondence value between two point clouds. In this paper, we employ the rejecting method based on Euclidean distance to eliminate invalid point pairs. Depending on the maximum detection range of

the radar and the detecting environment, we apply a threshold to point pairs and pairs whose distance is close or beyond the threshold will be considered invalid and rejected.

Normal vectors of measured points in this paper are utilized in error metric minimization. Using the nearest neighbors of each point, several fitting planes are estimated to calculate the normal vector of each point in the point cloud. By optimizing the target equation in (1) based on least square method, the normal vector of each point, which is perpendicular to vectors formed by the point and its neighbors, is obtained.

$$\boldsymbol{n} = \arg \min_{\boldsymbol{n}, \|\boldsymbol{n}\|=1} \sum_{i=1}^{k} ((\boldsymbol{p}_i - \boldsymbol{c})^{\mathrm{T}} \cdot \boldsymbol{n})^2 \tag{1}$$

where $\boldsymbol{n}$ represents the normal vector; $\boldsymbol{p}_i$ is the measured point in point cloud; k is the number of all nearest neighbors; $\boldsymbol{c}$ is the mean of k neighbors.

In experiment, we directly utilized singular value decomposition method to decompose matrix $\boldsymbol{M} = \sum_{i=1}^{K} (\boldsymbol{p}_i - \boldsymbol{c})$. Since the fitting plane represents a set of points with the same normal vector, meaning a great commonality, $\boldsymbol{n}$ equals the left singular vector with the minimal singular value.

After matching point pairs, the problem of calculating the optimal transformation matrix between two point clouds turns into minimizing the error metric in formula (2):

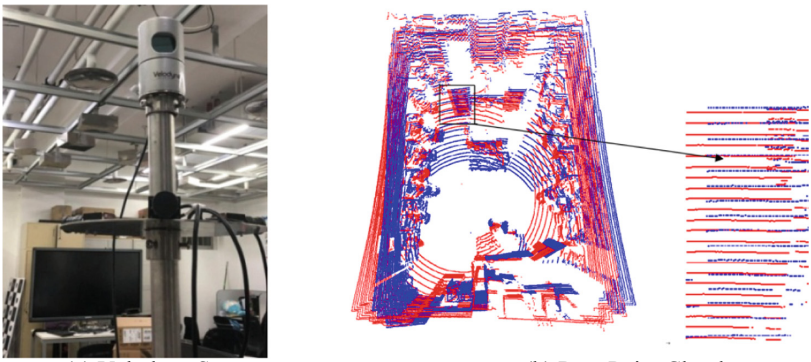
$$\boldsymbol{M}_{opt} = \arg \min_{\boldsymbol{M}} \sum_{i} ((\boldsymbol{M} \cdot \boldsymbol{p}_i - \boldsymbol{q}_i) \cdot \boldsymbol{n}_i)^2 \tag{2}$$

where $\boldsymbol{p}_i$ and $\boldsymbol{q}_i$ is the point in source point cloud, and the point in target point cloud, respectively; $\boldsymbol{n}_i$ is the normal vector of $\boldsymbol{q}_i$; $\boldsymbol{M}$ is the 4×4 transformation matrix.

Instead of solving the least square error metric by a standard non-linear method, (e.g. Levenberg-Marquardt method), which requires much computation time, we employed the linear least square method proposed by Low [4], where a liner least square optimization, approximate to the non-linear one, is applied to calculate transformation matrix $\boldsymbol{M}$ based on SVD. Once the distance between points is less than a certain threshold, the error metric is considered converged, with the final transformation value obtained.

3 Experiment

This experiment employed a Velodyne HDL-32E as the LiDAR equipment to sense the environment. The LiDAR can collect about 800,000 high-resolution 3D points per second. The developed algorithm was performed on an Inter® Core™ E3-1231 v3 CPU @3.40 GHz laptop computer 4 GB RAM. Figure 3(a) shows the equipment set up in our lab where the point cloud data was collected and Fig. 3(b) shows two raw point cloud data and a detailed view, collected in different frames, without implementing any registration method.



(a) Velodyne Sensor (b) Raw Point Cloud

Fig. 3. Experiment Equipment and Collected Point Cloud Data

A large number of ground points shown in Fig. 3(b) easily caused great disturbance to point pair matching, resulting in poor correspondence calculation. The selection and rejection step solved the matching dislocation between measured point clouds, achieving a more precise registration of two point clouds. Figure 4 shows the point clouds after the accurate registration based on point-to-plane ICP algorithm. With point selection and pair rejection, the registration result showed a higher accuracy and the detailed result is also shown in Fig. 4.

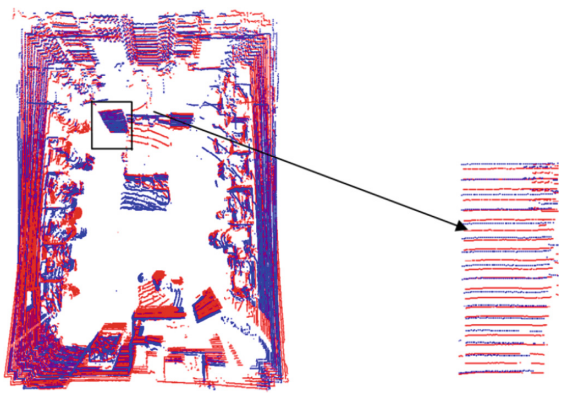


Fig. 4. Point Cloud after ICP-based Registration

The registration accuracy was calculated by the root mean square error (RMSE) of the distance between measured points. The experiment result of registering with and without point selection and rejection is compared in Table 1. The registration with point filtering showed higher accuracy and required fewer iteration times.

Table 1. Registration Result with and without Point Selection and Rejection

Registration without Point Selection and Rejection	Point Pair Number	69,120
	Iteration Number	21
	RMSE(m)	0.0192
Registration with Point Selection and Rejection	Point Pair Number	63,922
	Iteration Number	18
	RMSE(m)	0.0121

4 Conclusion

In order to achieve the precise registration of LiDAR point clouds for indoor environment modeling, this paper utilized an ICP-based registration method. Based on the point-to-plane ICP, we matched the nearest points between source and target cloud, and discarded invalid point pairs. After extracting the normal vectors of point data, the point-to-plane error metric was employed to iteratively calculate the optimal rotation and translation value between input point clouds. The experiment showed that two tested point clouds were precisely registered with our utilized method. Furthermore, since the point cloud sensed by LiDAR is sparse and non-structural, extracted normal vectors are not precise enough to provide an effective basis for subsequent registration, therefore, in the future work, we will adopt a more suitable solution to execute normal vector extraction, so as to achieve a more accurate registration of LiDAR point clouds.

Acknowledgments. This research was supported by National Natural Science Foundation of China (61503005), Beijing Natural Science Foundation (4184086), Beijing Young Topnotch Talents Cultivation Program (No. CIT&TCD201904009), NCUT "The Belt and Road" Talent Training Base Project, and NCUT "Yuyou" Project.

References

1. Boulch, A., Guerry, J., Saux, B.L.: SnapNet: 3D point cloud semantic labeling with 2D deep segmentation networks. Comput. Graph. **71**, 189–198 (2018)
2. Jaw, J.J., Chuang, T.Y.: Feature-based registration of terrestrial LiDAR point clouds. In: International Archives of Photogrammetry, ISPRS Commission III, vol. XXXVII, no. 3b, pp. 303–308 (2008)
3. Li, M., Wang, T., Li, Q.: Registration of 3D point cloud using a hierarchical object based method. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVIII, no. 5, Newcastle upon Tyne (2010)
4. Low, K.: Linear least-squares optimization for point-to-plane ICP surface registration. Technical report TR04–004, University of North Carolina, Chapel Hill (2004)
5. Besl, P.J.: McKay: a method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. **14**, 239–256 (1992)
6. Chen, Y., Medioni, G.: Object modeling by registration of multiple range images. Image Vis. Comput. **10**, 145–155 (1992)

Particle Color Transforming Method by Mapping Image Texture Arrangement with Weight Value

Jia-Yi Qiu(✉), Fei Hao, and Jae-Hyuk Ko

China and South Korea School of New Media, Zhongnan University of Economics and Law, Wuhan 430073, People's Republic of China
550139284@qq.com

Abstract. The present research relates to a method for transforming a particle color based on a weighted image texture mapping. More particularly, to a particle color mapping method in a Maya graphical tool, particle color mapping can be performed without using expressions. As the color mapping is performed by the image texture, any type of image can be represented using particles. In addition, since particles are grouped and color-mapped, to a weighted image texture mapping-based particle color conversion method that can be augmented.

Keywords: Maya · Particle color · Image texture · Weight value

1 Introduction

1.1 Research Background

Particles refer to the technique of simulating specific phenomena in computer graphics. Its advantages are to simulate some specific images of natural phenomena, physical phenomena and spatial distortions, which is convenient for us to realize some real natural and random images. Various image/effect implementations using particles have been performed in various ways in computer graphics work. In Maya, which is widely used as a computer graphics tool, it is impossible to directly modify the position value of a particle, or to connect an external input with a position value of a particle, or to connect an external input and a position value. There is no limit to the color mapping that can be done only through the particle representation provided by Maya.

On the other hand, a color is set and assigned to an image texture array mapped to a specific space in the virtual space of the CG image generated by the graphic tool Maya, and when the particles moving in the virtual space of the CG image pass through the image texture array, So that the color assigned to the particle is applied to the particle. In such image texture-based particle color mapping, it is inconvenient to map all the large particles to virtual space, and then apply color to each of the particles corresponding to the unit area of the image texture array.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 345–350, 2020.
https://doi.org/10.1007/978-981-32-9244-4_49

1.2 Research Purposes

Accordingly, the present research has been made to solve the problems of the related art, and it is an object of the present research to provide an image texture array that is mapped to a specific space in a virtual space of a CG image generated by Maya, A weight value is assigned to a region of a size. When a moving particle in a virtual image space of a CG image passes through an image texture array, a weighted color is assigned to the image texture array, So that the particle color mapping can be performed without using the particle expressions provided in Maya, so that any type of image can be represented using the particles, and a new type weighted image texture mapping based particle color conversion method in which the efficiency of arithmetic can be increased by grouping and color mapping the particles.

2 Research Method

2.1 Research Outline

According to an aspect of the present research, there is provided a method of operating a graphical tool, A particle generation step in which a particle aggregate composed of a plurality of particles is generated by the Maya and output to the virtual space of the CG image; An image texture generation step in which the image texture is generated by the Maya; An image texture array conversion step of converting an image texture into an image texture array in which a unit area of a set size is arranged in a plurality of rows and columns; An image texture array weighting step of assigning weight values to areas of a set size belonging to an image texture array; An image texture array position value corresponding color setting step wherein a color corresponding to a position value of each unit area constituting the image texture array is set by the Maya; An image texture array mapping step in which an image texture array is mapped to a specific space inside the virtual space of the CG image; Particles are placed in the interior region of the image texture array by movement of the particle aggregate and the color assigned to the unit region of the image texture array corresponding to the current position of the particle is assigned to the particle, And a particle color conversion step in which a color to which the weight given to the area is applied is assigned to the particle.

In the weighted image texture mapping-based particle color conversion method according to the present research, the image texture array conversion step may include a unit area constituting the image texture array composed of one or more pixels.

In the weighted image texture mapping based particle color conversion method according to the present research, the image texture array weighting step may include one or more unit areas having a set size to which weights are assigned.

In the weighted image texture mapping based particle color conversion method according to the present research, in the color matching step, the color corresponding to each unit area position value of the image texture array is separated into RGB color and alpha.

In the weighted image texture mapping based particle color conversion method according to the present research, the particle color conversion step may include: a

particle-image texture array associating step in which particles are arranged in an image texture array array while inducing movement of the particle aggregate; An image texture array position value connection step of connecting a position value of a particle belonging to a unit area constituting an image texture array and a position value of the unit area to each other; A particle attribute changing step in which an attribute transfer imparted to a particle to which a unit area and a position value of the image texture array are connected is changed to a color attribute; The color assigned to the unit area position value of the image texture array is assigned to the attribute of the particle connected to the unit area, so that color mapping is performed, and the color to which the weight assigned to the area to which the unit area of the image texture array belongs, a particle color mapping step of assigning a particle color; And a color conversion completion step of converting the color of the particle by applying a color separated by RGB color and alpha.

2.2 Detailed Description of the Research

Hereinafter, embodiments of the present research will be described in detail with reference to the accompanying drawings. In the drawings and the detailed description, there are shown general computer graphics, graphics tools, Maya, particles, particle creation, particle color mapping, attribute transfer, image texture, image texture arrangement, pixel, The description and the operation of the structure and operation that the workers of the present research can easily understand are omitted or simplified. In particular, in the drawings and the detailed description, detailed descriptions and illustrations of the technical structures and functions of the elements not directly related to the technical features of the present research are omitted and only the technical structures related to the present research are shown or described briefly respectively.

The weighted image texture mapping based particle color transformation method according to an embodiment of the present research includes a graphic tool activation step, a particle creation step, an image texture creation step, an image texture array transformation step, an image texture arrangement weighting step, an image texture array position value correspondence color setting step, an image texture array mapping step, and a particle color conversion step.

The step of activating the graphic tool is the step of activating Maya which is a graphic tool (20) installed in the computer device (10) as shown in Fig. 1.

The particle generation step is a step in which the particle aggregate (2) composed of a plurality of particles (1) is generated by Maya and output to the virtual space (11) of the CG image. Such a particle creation step can generate a particle by generating a particle grid.

The image texture generation step is the stage in which the image texture (3) is generated by Maya. The image texture can be composed of various shapes and various color combinations, and can be composed of various images.

The image texture array conversion step is a step of converting the image texture (3) into the image texture array (4) in which the unit area (5) of the set size is arranged in a plurality of rows and columns. In this image texture array conversion step, the unit area (5) constituting the image texture array (4) is made up of one or more pixels.

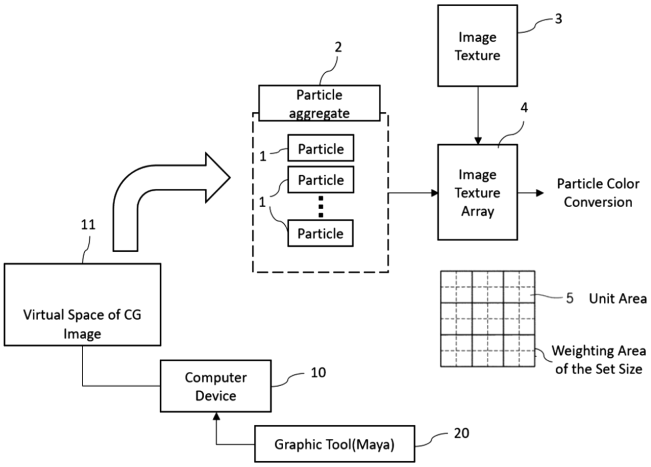


Fig. 1. A device configuration diagram for implementing a weighted image texture mapping based particle color conversion method according to an embodiment of the present research.

The image texture array weighting step is a step in which a weight value is assigned to each area of the set size belonging to the image texture array (4). Here, the area of the set size to which the weight is assigned may be one or more unit areas (5). The weight may have a value of 0 and 1, a value of 0 to 1, or a value of 1 or more.

The image texture array position value corresponding color setting step is a step in which the color corresponding to the position value of each unit area (5) constituting the image texture array (4) is set by Maya. In the image texture array position value correspondence color setting step, the color assigned to the position value of each unit area (5) has an attribute of a point color.

Here, the image texture array position value corresponding color setting step according to the embodiment of the present research allows the color corresponding to the position value of each unit area (5) of the image texture array (4) to be set separately from RGB color and alpha. This is because Maya prefers RGB.

The image texture array mapping step is a step in which the image texture array (4) is mapped to a specific space inside the virtual space (11) of the CG image.

In the particle color conversion step, the particles (1) are arranged in the inner area of the image texture array (4) by the movement of the particle aggregate (2), and the unit area of the image texture array (4) corresponding to the current position of the particle (1) The color assigned to the particle (1) is imparted to the particle (1) and the movement of the particle (1) to which the attribute transfer is applied causes the particle (1) to be present in the area inside the image texture array (4) And the color of the particle (1) is converted while the color assigned to the position value of the image texture array corresponding to the particle position value is assigned to the particle (1). In the particle color conversion step according to the present research, the color to which the weight assigned to the area to which the unit area (5) of the image texture array (4) belongs is applied to the particle (1).

For this purpose, the particle color conversion step according to an embodiment of the present research includes a particle-image texture array association step, a particle-image texture array position value connection step, a particle attribute change step, a particle color mapping step, color conversion completion step (Fig. 2).

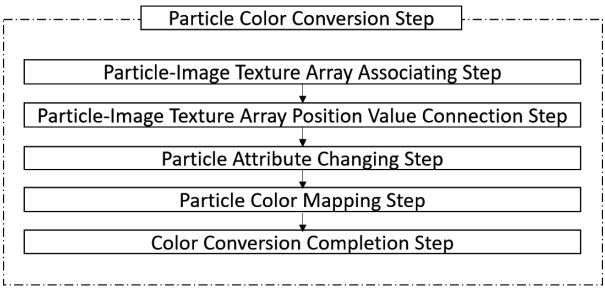


Fig. 2. Detailed order block diagram of the particle color conversion step in the weighted image texture mapping based particle color conversion method according to the embodiment of the present study.

The particle-image texture array associating step is a step in which the particle (1) is placed in the area inside the image texture array (4) while the movement of the particle aggregate (2) is induced. The movement of the particle aggregate (2) can be performed by attribute transfer given to the particle (1).

The particle-image texture array position value connecting step is a step in which the position value of the particle (1) belonging to the unit area (5) constituting the image texture array (4) and the position value of the unit area (5) are connected to each other.

The particle attribute changing step is a step in which the attribute transfer given to the particle (1) to which the unit area (5) of the image texture array (4) and the positional value are connected is changed to a color attribute.

The particle color mapping step is a step in which the color assigned to the position value of the unit area (5) of the image texture array (4) is assigned to the attribute of the particle (1) connected to the unit area (5) to perform color mapping. Here, the particle color mapping step according to the present research allows the color to which the weight assigned to the area to which the unit area (5) of the image texture array (4) belongs is assigned to the particle (1). The color conversion completion step is a step in which the color of the particle (1) is converted by the application of the RGB color and the color separated by the alpha. The weighted image texture mapping based particle color conversion method according to an embodiment of the present research configured as described above is a method of transforming an image texture array that is mapped to a specific space in a virtual space (11) of a CG image generated by a Maya, A weight is assigned to an area of a set size for dividing the image texture array (4) and a particle (1) moving in the virtual space (11) of the CG image is assigned a weight value. The color assigned to the image texture array (4) at the time of passing through the image texture array (4) and weighted is applied to the particle (1) so that the

particle color mapping is performed so that the particle color mapping can be performed through the particle expression fundamentally provided in Maya Particle color mapping can be performed without any type of image being rendered using particles, and particles can be grouped into color So that the computation efficiency can be increased. Figure 3 is an image published using this technology at an international exhibition. (2nd IACC at FRESCO in TEXAS, USA)

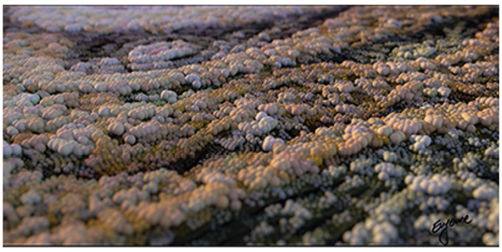


Fig. 3. Published images at 2nd IACC at FRESCO in TEXAS, USA.

3 Conclusion

In Maya, which is widely used as a computer graphics tool, it is impossible to directly modify the position of a particle or to link an external input with a position value of a particle, or to link an external input and a position value. There is a limit to the ability to do color mapping only through the particle expressions that are provided by Maya only. According to the weighted image texture mapping based particle color transformation method of this study, the particle color mapping is done by linking the image texture array with the weighted value of the particle and the particle color mapping is done without going through the underlying particle expression in Maya, In particular, according to the weighted image texture mapping based particle color conversion method according to the present research, particles are grouped and color-mapped to increase the computation efficiency.

References

1. Hai-Wen, Y.: Maya particle rendering properties. Comput. Knowl. Technol. **9**(30), 6899–6901 (2013)
2. Wang, S.-W., Chen, X.-Q., Li, S., Liu, P.: Film and television special effects particle system research. Softw. Guide **11**(04), 168–170 (2012)
3. Wang, X.-X.: Application of Maya particle power system in film and TV special effects. JiLin University of the Arts (2016)

A Customer Group Mining Method Based on Cluster Analysis

Yongping Tang(✉) and Zizhen Peng

Department of Mechanical and Electrical Engineering, WuZhou Vocational College, WuZhou 543002, China
tyongping@163.com, pengzizhen@qq.com

Abstract. With the rapid development of WeChat's business economy, customer data is exploding. Taking the tortoise herb jelly WeChat marketing data in WuZhou as an example, how to accurately analyze customer data, discover customers' consumption characteristics of the tortoise herb jelly, and provide different efficacy tortoise herb jelly products pertinently to different customer groups, has become a major problem in the development of enterprises. In order to solve the above problems, this paper applies clustering analysis algorithm to the process of customer group mining, uses the key information of existing customers to classify customer categories, and obtain higher latitude customer information. Experiments show that customer segmentation based on clustering results can improve the analysis efficiency of the relationship between customer groups and products, identify important factors affecting product sales, and provide support basis for enterprises to carry out customer-centered precision service.

Keywords: WeChat business · Data mining · Customer group · Cluster analysis

1 Introduction

With the rapid development of WeChat business economy [1], WeChat marketing has become the main way to sell tortoise herb jelly [2]. At the same time that customers are growing explosively, enterprises also collect a large amount of transaction data through the marketing platform [3, 4]. Taking the tortoise herb jelly WeChat marketing data in WuZhou as an example, how to accurately analyze customer data, discover customers' consumption characteristics of the tortoise herb jelly, and provide different efficacy tortoise herb jelly products pertinently to different customer groups, has become a major issue in the development of enterprises. In order to solve this problem, this paper takes the mass data of WeChat marketing of tortoise herb jelly as the research background, constructs structured customer analysis data set by extracting multi-source data related to tortoise herb jelly consumption, proposes a method of customer group mining

This work was supported by the Basic Ability Promotion Project for Young and Middle-aged Teachers in Universities in Guangxi province of China (Grant NO.2018KY1236).

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 351–357, 2020.
https://doi.org/10.1007/978-981-32-9244-4_50

based on mass data, and establishes PEMST data model and K-Means clustering model to mine and analyze customer data set. On the one hand, the model can explore the characteristics of different customer groups, analyze the relationship between different customer groups and a certain type of tortoise herb jelly products, and provide customers with accurate products based on the relevance in the process of customer consumption; on the other hand, it can help companies find out the main factors which impact product sales, targeted development of marketable products, improve the accuracy of marketing, support enterprises to carry out customer-centric precision services. Experiments show that the clustering model based on K-Means algorithm can effectively mine a large number of customer groups and improve the efficiency of data analysis. It is suitable for large-scale data analysis of enterprises to find valuable customer groups.

2 Problem Statement

In the face of fierce market competition, WuZhou tortoise herb jelly Enterprise is also facing business crisis such as customer loss, competitive decline and inaccurate product positioning. The advent of the information age has transformed the focus of business operations from product center to customer center. How to find high-value customer groups has gradually become a core issue for enterprises. This paper obtains the analysis results needed to solve the above problems by constructing a reasonable customer value assessment model and conducting data mining.

The customer value model mainly includes five indicators: age, sex, product_type (product type), lottery_count (lottery count), and province (purchase province). This paper identifies customer groups with different values through customer segmentation, and built the PEMST model.

According to the above indicators, how to mine high-value customer groups from the PEMST model can be formalized as:

$$\begin{cases} M := \arg \max\limits_{v} f(t, v) \\ N := \arg \min\limits_{c} g(t, c) \end{cases} \tag{1}$$

where t is the customer type, v is the business value, and $f()$ is the relationship function between the customer type and the business value. From the perspective of the enterprise, the business value M obtained by selecting the appropriate customer type t is the larger the better; c is the operating cost of the enterprise, and the operating cost N of the enterprise selected by the appropriate customer type t is the lower the better.

3 Construction of Mining Models

3.1 Data Extraction and Preprocessing

Due to the large number of attributes in the original data, this paper selects only five attributes related to the PEMST indicator according to the customer value PEMST model of the enterprise WeChat platform: E = age, S = sex, T = product_type (product type), M = lottery_count (lottery count), P = province (purchase province).

The corresponding product types are: *1* is the original tortoise herb jelly; *2* is the beauty acne tortoise herb jelly; *3* is the lotus seed & Job's tears tortoise herb jelly; *4* is the conditioning caring nervous tortoise herb jelly; *5* is the LuoHan chrysanthemum tortoise herb jelly.

After data cleaning by attributes construction and data standardization, this paper obtains a format that meets the needs of mining tasks and algorithms, as shown in Table 1.

Table 1. Data set after property selection

Customer name	age	sex	product_type	lottery_count	province
dodo	21	1	5	2	GuangXi
Emma	22	1	2	3	GuangDong
Abby	16	2	1	2	GuangXi
…					

3.2 Model Building

In order to solve the problem defined by Eq. (1), we proposed the customer value analysis model, and its construction process is mainly divided into two parts: firstly we proposed the PEMST data model, secondly we used the K-Means clustering algorithm to cluster the customers.

K-Means is one of the clustering algorithms, where K represents the number of categories and Means represents the mean value. As the name implies, K-Means is an algorithm that clusters data points by mean value [5, 6]. The K-Means algorithm divides similar data points by the preset K value and the initial centroid of each category, and obtains the optimal clustering result through the divided mean iterative optimization.

In order to analyze high-value customers from the customer group and their purchasing power data, we need to determine the number of initial centroids and classify them according to the distance between the group and its purchasing power to the centroid. We use the Euclidean distance to calculate the similarity, the smaller the distance, the higher the similarity between the two users. We assume that the sample data set is $D = \{x_1, x_2, \cdots, x_u, \cdots, x_n\}$, then the data of the *u-th* user is $x_u = \left(x_u^i, x_u^j\right)$, where x_u^i is the purchase quantity of user u, and x_u^j is the group attribute of user u. Let the *k-th* initial centroid is $\mu_k = \left(\mu_k^i, \mu_k^j\right)$. Then the similarity d of x_u and μ_k is calculated as:

$$d = \|x_u - \mu_k\|_2 = \sqrt{\left(x_u^i - \mu_k^i\right)^2 + \left(x_u^j - \mu_k^j\right)^2} \tag{2}$$

The K-Means algorithm mainly divides the sample set into K clusters according to the distance between the given sample sets (that is the distance d between x_u and μ_k above), make the distance between the clusters is as large as possible and the points distance between each other in the cluster is as small as possible. Specifically, it can be refined into the following steps:

Step 1: Select K objects in the data space as the initial center, and each object represents a cluster center;

$$\mu_k = \frac{1}{|c_k|} \sum_{x_u \in c_k} x_u \tag{3}$$

where $|C_k|$ represents the number of data objects in the *k-th* class cluster, the summation in this formula refers to the sum of all the elements in the class cluster C_k on each column attribute, so C_k is also a vector containing E attributes, expressed as:

$$C_k = (C_{k,1}, C_{k,2}, \ldots C_{k,E}) \tag{4}$$

Step 2: According to the Euclidean distance between the data objects in the sample and the clustering centers, they are allocated to the nearest clustering (the highest similarity) according to the nearest criterion.

Step 3: Update the cluster center. Calculating the value of the objective function by using the mean value corresponding to all objects in each category as the cluster center of the category;

Step 4: Judge whether the values of the cluster center and the objective function change or not. If they do not change, then output result. If they change, return to step *2* and repeat the process. The specific process is shown in the following pseudocode:

Algorithm 1 A Customer Mining Algorithm Based on K-Means

Require: The sample data set $D = \{x_1, x_2, \cdots, x_u, \cdots, x_n\}$, number of clusters k, threshold α.

Ensure:

1: Initialize: $k=5$, mean vector $\{\mu_1, \mu_2, \cdots, \mu_k\}$.
2: **repeat**
3: for j in 1:m do
4: Calculate the similarity d according to *Eq. (2)*
5: for i in 1:k do
6: Calculate $\mu_k = \frac{1}{|c_k|} \sum_{x_u \in c_k} x_u$
7: util: $\left(VE = \sum_i^k d\right) \le \alpha$

Output: Cluster division $C_k = \left(C_{k,1}, C_{k,2}, \cdots, C_{k,E}\right)$

As with the above principle, we perform clustering calculation by combining the customer group with the product type, the product type with the regional distribution, and finally get the best customer group classification.

4 Experiment

4.1 Analyze the Relationship Between Customer Groups and Product Types

Using the data model we have established to analyze the relationship between customer groups and commodity types, as shown in Figs. 1 and 2. Group *1* (under 30 years old customer group) has the highest consumption ratio of No. *2* (beauty acne tortoise herb jelly); Group 4's (over 60 years old customer group) favorite product is No. *4* (conditioning &caring nervous tortoise herb jelly); Group *2* (customer group between *30* and *45* years old) and Group *3* (customer group between *45* and *59* years old) prefer the No. *1* (original tortoise herb jelly). According to the above analysis results, enterprises can clearly know which are the high-consumer groups, and which are the middle and low consumer groups, and according to the consumption behavior of different customer groups, accurately develop and produce marketable products for them, and improve the competition of enterprises.

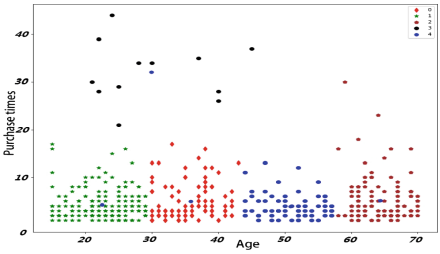


Fig. 1. Analysis diagram of the relationship between customer groups and commodity types

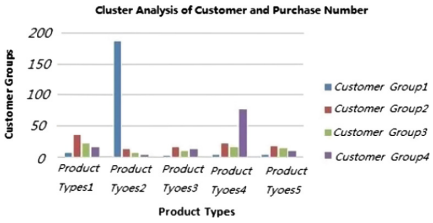


Fig. 2. Cluster analysis of customer and purchase number

4.2 Analyze the Relationship Between Product Type and Regional Distribution

Through cluster analysis, we can accurately find out where Wuzhou tortoise herb jelly has a high sales volume. From Figs. 3 and 4, we can see that the largest sales areas of tortoise herb jelly products are Guangxi and Guangdong, and different customer groups have different needs for tortoise herb jelly. Therefore, enterprises can develop new products according to the analysis results to meet the needs of different customer groups. At the same time, enterprises can maintain different customer groups accurately according to the analysis results, improve customer loyalty to enterprise products, and achieve stable and rapid development of enterprises.

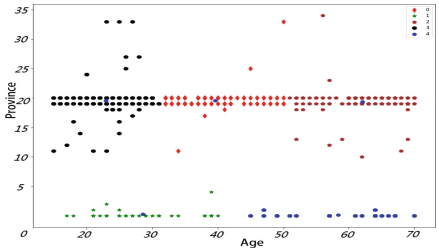


Fig. 3. Analysis of purchasetime and purchase area

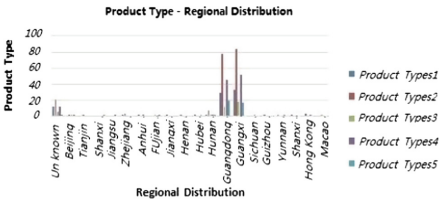


Fig. 4. Product type - regional distribution

Based on the above analysis results, we define four customer categories: the most valuable customer; the important value customer; the general value customer; the low value customer (1). According to the characteristics of customer groups, we rank all kinds of customer groups, and the results are shown in Table 2.

Table 2. Customer group value ranking

Customer groups	Commodity type	Ranking	Category
Group 1	2	1	the most valuable customer
Group 4	4	2	the important value customer
Group 2	1	3	the general value customer
Group 3	5	4	the low value customer

5 Conclusion and Future Work

In order to enable enterprises to mine customer behavior from complex data and analyze customer preferences, guide product production and support the formulation of marketing strategies. We propose a clustering-based customer group mining method, which includes PEMST data model and K-means customer clustering model.

The experimental results show that enterprises can use our model and algorithm to analyze customer data. The results can help enterprises better distinguish customer groups with different values, obtain the distribution of commodity types and regions, and the preferences of different customer groups for commodity types, thus reducing customer maintenance costs and improving business efficiency. The model constructed in this paper can also be applied to other industries or enterprises that need customer maintenance to facilitate their differentiated management of different customer groups, so as to make customer service more precise and improve the competitiveness of enterprises.

References

1. Wang, G., Zhang, W., Zeng, R.: WeChat use intensity and social support: the moderating effect of motivators for WeChat use. Comput. Hum. Behav. **91**, 244–251 (2019)
2. Baidu, Olecranon turtle and smilax glabra do with Traditional Chinese medicine. https://baike.baidu.com/item/%E6%A2%A7%E5%B7%9E%E9%BE%9F%E8%8B%93%E8%86%8F/3313726
3. China Internet Network Information Center, Statistical Report on the Development of Internet in China, http://www.cac.gov.cn/2018-08/20/c_1123296882.htm
4. Jing, R.: On modeling and maximizing online customer group value. DEStech Trans. Comput. Sci. Eng. (icmsa) 2018
5. Nguyen, C.D., Duong, T.H.: K-means**–a fast and efficient K-means algorithms. Int. J. Intell. Inf. Database Syst. **11**(1), 27–45 (2018)
6. Lakshmanaprabu, S.K., Shankar, K., Gupta, D., Khanna, A., Rodrigues, J.J.P.C., Pinheiro, P. R., de Albuquerque, V.H.C.: Ranking analysis for online customer reviews of products using opinion mining with clustering. Complexity (2018)

HPFLRF: A High Performance Fingerprint Localization Algorithm Based on Random Forest

Pengyu Huang[1(✉)], Haojie Zhao[1], and Wei Wang[2]

[1] School of Telecommunication and Engineering, State Key Labs of ISN, Xidian University, Xian, Shaanxi, People's Republic of China
pyhuang@mail.xidian.edu.cn, hjzhao@stu.xidian.edu.cn

[2] School of Information Science and Technology, Northwest University, Xian, Shaanxi, People's Republic of China
wwang@nwu.edu.cn

Abstract. In recent years, indoor localization base on fingerprint has become more common. Due to the complexity and variability of indoor environment, it is difficult for traditional indoor localization algorithm to obtain better localization accuracy and stability. In this paper, we propose a high performance fingerprint localization algorithm based on random forest (HPFLRF), which has higher precision and stability. Our algorithm could select a valid subset of APs through multiple AP selection method. In addition, our algorithm uses the random forest training positioning model to improve the stability of the algorithm effectively, and overcome the problem of overfitting in single decision tree model. The results of experiment show that our algorithm has better localization performance which average positioning error is 1.3718 m, only one seventh of the localization algorithms based on multiple times AP selection and decision tree.

Keywords: AP selection · Random forest · Positioning stability · Indoor localization

1 Introduction

With the increasing popularity of smart mobile devices, users have higher demand for location-based services (LBS) include high precision indoor positioning. In recent years, many researchers have focus on the research of indoor positioning algorithm. Although the outdoor positioning based on GPS has been able to achieve very high positioning accuracy. However, due to the complexity and diversity of indoor

P. Huang—The financial support of the program of Key Industry Innovation Chain of Shaanxi Province, China (2017ZDCXL-GY-04-02), of the program of Xi'an Science and Technology Plan (201805029YD7CG13(5)), Shaanxi, China, of Key R&D Program – The Industry Project of Shaanxi (Grant No. 2018GY-017), of Key R&D Program – The Industry Project of Shaanxi (Grant No. 2017GY-191) and of Education Department of Shaanxi Province Natural Science Foundation, China (15JK1742) are gratefully acknowledged.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 358–363, 2020.
https://doi.org/10.1007/978-981-32-9244-4_51

environment, there are multipath effects, attenuation, diffraction and scattering phenomena in indoor signal transmission, so it is more difficult to obtain high precision indoor position coordination.

In recent years, a large number of indoor localization algorithms have been proposed as solutions for accurate indoor localization. In [1], an AP selection method based on information gain was proposed to improve the effectiveness of AP selection, and location based on the decision tree model. In [2], FS-kNN is proposed to improve localization accuracy. They build up a novel RSS-level based feature scaling model which is used to improve positioning accuracy. Paper [3] proposed a positioning algorithm based on neural network modeling. The algorithm can get a better model of Wi-Fi signal, and improve positioning accuracy. In [4], the author proposes an original algorithm based on machine learning. Moreover, this paper investigates several kernel-based machine-learning techniques, such as LS SVM, SVR. Moreover, the DeepFi is proposed in paper [5], which is a novel deep learning method based on indoor fingerprinting system with Channel State Information (CSI). This paper use a greedy learning algorithm to reduce complexity by training the weights layer-by-layer, and use a probabilistic method based on radial basis function to obtain the estimated location.

The original random forest algorithm is proposed in [6], which is a supervised machine-learning algorithm. Each random forest model is composed of multiple decision tree models. Random forest algorithm has strong generalization ability and can prevent the problem of overfitting. The algorithm of this paper construct a positioning model based on random forest algorithm with good positioning performance. In addition, HPFLRF selects valid APs by multiple AP selection method, which can further improve the precision of location.

The rest of the paper is organized as follows: Sect. 2 describes HPFLRF in detail. Section 3 is about the experimental results and analysis. Section 4 is a conclusion of MTAPS.

2 Algorithm Description

HPFLRF consists of two parts: offline phase and online phase. There are four steps in the offline phase, AP data collection, AP selection, location clustering and random forest location model construction.

2.1 AP Data Collection

The purpose of this step is to build up the initial fingerprint database by collecting AP's RSSI data in the localization environment. The detailed process is as follows. Firstly, we divide the positioning environment into squares of the same size. The center of each square is marked as the reference position of the square. Each reference position has a label and a two-dimensional coordinate. At each reference position, collecting AP's RSSI data for a period. Finally, we use these AP data to form the initial fingerprint database.

2.2 AP Selection

Currently, WiFi access points are popularity in our living environment, so a large number of APs can be detected in the location environment. At the same time, this can lead to dimension increasing of the fingerprint database. In addition, due to the complexity of indoor environment, some APs in the detected APs set have bad effect on positioning accuracy. Therefore, it is necessary to select suitable subset of APs from all APs. This selection step can effectively improve the performance of positioning algorithm. Therefore, as in this paper, HPFLRF runs AP selection method twice to improve the performance of location. The process is as follows:

(1) Obtain primary AP subset

First, set *th1* as the signal strength threshold of AP. Then, calculating the average signal strength of each AP detected at each location during the collection time. Finally, comparing the average signal strength of each AP with *th1*. Abandon the APs that average signal strength is less than *th1*, and merge the remaining APs into primary AP subset Pr*iAP*.

(2) Obtain fingerprint AP subset

Information gain (InfoGain) algorithm [1] is proposed as a new AP selection method based on information theory. InfoGain algorithm uses AP's information gain to represent the AP's discriminative power to location. The more discriminative power, the more important AP is. In this sub-step, our algorithm calculates the information gain of each AP in Pr*iAP* by using InfoGain algorithm. According to the information gain, all APs are sorted in descending order, and then the first k APs with the larger information gain are selected to constitute the subset of fingerprint AP Finger*AP*. Then structuring fingerprint vector of all reference locations based on Finger*AP* and forming the location fingerprint database based on fingerprint vector.

2.3 Clustering Location and Making up Random Forest Decision Model

The traditional fingerprint location algorithms usually matches the locating sample with the fingerprint of all reference positions, and chose the position with the highest matching degree as the target position. As in our algorithm, we divides all the reference positions into clusters and places the reference positions with high similarity into a group. In this way, the real-time location data only needs to match with a part of reference locations, which can effectively improve the efficiency of the algorithm. In our algorithm, we use a classical cluster algorithm, k-means algorithm [7], to cluster reference locations. When the location clustering is finished, we can get the fingerprint set of each location cluster.

Then, HPFLRF builds up a random forest decision model for each location cluster. The process is as follows: build up random forest model for cluster i, and cluster i contains *m* reference locations. The sampled data at these reference locations constitute the original training sample set OriSet of cluster i, and the size of sample set is N.

(1) Determine the number of decision tree that is contained in each random forest M and the number of attributes K that is used in decision tree, which is the number of APs.
(2) Based on the idea of set-back sampling, M training data sets are randomly extracted from OriSet, and the size of each training set is N.
(3) Randomly generate M AP sets based on Pr *iAP*, and the size of each AP set is K;
(4) Map the M training data sets to M AP sets one by one, and use them to train the decision tree models for every random forest. Each decision tree model is trained by C4.5 algorithm [8].

2.4 Online Phase

As in online phase, HPFLRF includes two steps. The process is as follows:

(1) Determine the real-time location data belongs to which cluster. Calculating the Euclidean distance between location data and each location cluster, and select the location cluster as the target cluster, which has the smallest Euclidean distance to the location data.
(2) Decide the target location of real-time location data. In previous sub-step, HPFLRF get the cluster that real-time location data belongs to. In this sub-step, HPFLRF uses the random forest model of this cluster to obtain the target location. First, each random forest contains M decision tree models, and each decision tree model get a positioning result for the location data. Therefore, HPFLRF can form a positioning result set through random forest, which size is M. Then, counting the times of each result in the result set and sort the positioning results in descending order. If there is only one result with the largest proportion, this result is the final target result. Otherwise, if there are multiple results with the largest proportion, the average value of these positioning result coordinates is taken as the final target result.

3 Experimental Evaluation

In this section, we evaluate the performance of HPFLRF through several experiments. As Fig. 1 shows, the experiment is carried out in the fourth-floor corridor of main building of Xidian University. The experimental testbed include 177 reference locations, and every reference location is a grid of 0.8 m * 0.8 m. In the phase of collecting AP data, we collect 50 samples at every reference location. Each sample lasts six seconds. As a result, we can scan more than one hundred APs at every reference location in our localization environment.

We compare the positioning performance of HPFLRF with the positioning algorithm based on InfoGain algorithm and C4.5 decision tree, as well as the positioning algorithm based on multiple times AP selection and C4.5 decision tree. Figure 2 shows the 20 positioning test results of 33 positioning samples. It can be seen from the figure

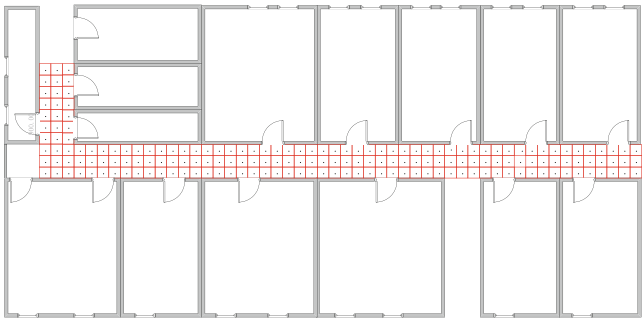


Fig. 1. Experimental testbed

that HPFLRF has good positioning accuracy and smaller fluctuation of positioning result. The average of HPFLRF is 1.3607 m, and the variance of the average positioning error is 0.0171. The average positioning error of the other two algorithms is 1.5152 m and 1.7465 m.

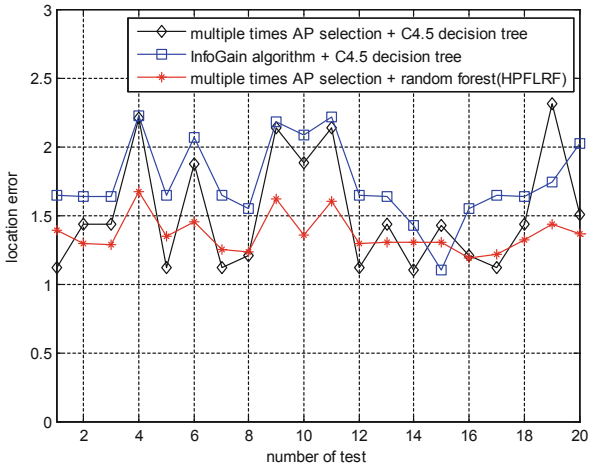


Fig. 2. Comparison of positioning performance of the three positioning algorithm

Figure 3 shows the influence of different clustering numbers on the three algorithms. From the figure, we can see that HPFLRF has the best positioning accuracy under the condition of the same number of clusters. And when the number of clusters is 1, that is, no location clustering is conducted, the positioning accuracy of HPFLRF is the worst. Therefore, location clustering can improve positioning accuracy.

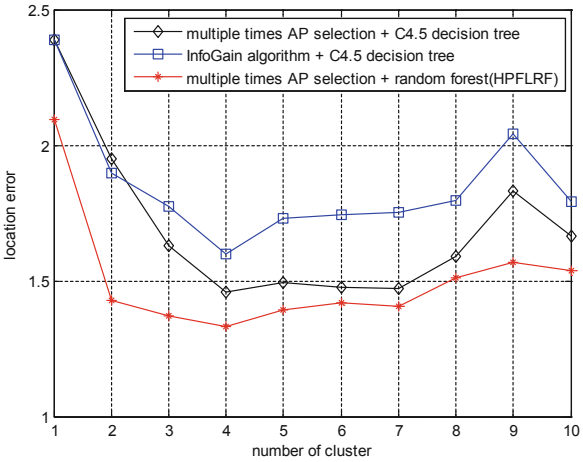


Fig. 3. The influence of different clustering numbers on the three algorithms

4 Conclusion

This paper proposes a high performance fingerprint localization algorithm based on random forest. HPFLRF could improve the performance of AP selection algorithm by multiple times AP selection. In addition, HPFLRF based on the random forest theory to establish positioning model, which can effectively promote positioning accuracy, and improve the stability of the algorithm. HPFLRF could overcome the problem of overfitting effectively by random forest theory, and does not need any pruning. The results of the experiments show that our algorithm has good positioning performance.

References

1. Chen, Y., Yang, Q., Yin, J., Chai, X.: Power-efficient access-point selection for indoor location estimation. IEEE Trans. Knowl. Data Eng. **18**(7), 877–888 (2006)
2. Li, D., Zhang, B., Yao, Z., Li, C.: A feature scaling based k-nearest neighbor algorithm for indoor positioning system. In: 2014 IEEE Global Communications Conference, Austin, TX, pp. 436-441 (2014)
3. Mok, E., Cheung, B.K.-S.: An improved neural network training algorithm for Wi-Fi fingerprinting positioning. ISPRS Int. J. Geo-Inf. **2**(3), 854–868 (2013)
4. Mahfouz, S., Mourad-Chehade, F., Honeine, P., Farah, J., Snoussi, H.: Kernel-based machine learning using radio-fingerprints for localization in WSNs. IEEE Trans. Aerosp. Electron. Syst. **51**(2), 1324–1336 (2015)
5. Wang, X., Gao, L., Mao, S., Pandey, S.: CSI-based fingerprinting for indoor localization: a deep learning approach. IEEE Trans. Veh. Technol. **66**(1), 763–776 (2017)
6. Breiman, L.: Machine Learning **45**, 5 (2001)
7. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification, 2nd edn. Wiley-Interscience, New York (2000)
8. Quinlan, J.R.: C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers Inc., San Francisco (1993)

A Study of an Unmanned Surface Vehicle System for Marine Environmental Monitoring

NamHyun Yoo(✉)

Department of Computer Science and Engineering, KyungNam University, 7 Kyungnamdaehak-ro, Masanhappo-gu, Changwon-si, GyeongNam 51767, Republic of Korea
hyun43@kyungnam.ac.kr

Abstract. Recently, drones have been widely used in various fields. The most active field is aviation drones, where various drones are used in combat to conduct unmanned surveillance and attacks, and in recent years the US military has trained drones pilots. In addition, it is used in various fields such as forest fire monitoring, border surveillance, broadcasting, aerial surveys, environmental monitoring, and courier delivery. In recent years, the range of applications for personal use and leisure has been widening. The next most frequently used field is the automotive sector. In the case of vehicles, autonomous navigation is added to general automobiles, and unmanned vehicles are expected to be realized within 10 years. On the other hand, the use of drones in the marine sector is somewhat lower than in aviation or the field of vehicles. In particular, many companies are conducting research and development in the aviation and automobile fields, but only a few representative companies are conducting research and development in the marine field. Conventional marine drones are mostly utilized in the public and military fields, but MR-SENTINEL proposed in this paper is a system used for private sector such as aqua farming and marine environmental monitoring. The MR-SENTINEL designed and developed in this paper uses an electric propulsion motor with a length of 3 m. MR-SENTINEL's communication system basically uses WiFi and enables free operation without additional communication fee. In addition, existing systems used expensive sensors, but MR-SENTINEL uses low-cost sensors to lower the price, so it can be widely used in the private sector.

Keywords: USV · Remote operation · Marine drone · Electric propulsion · ROS

1 Introduction

Recently, drones have been widely used in various fields. The most active field is aviation drones, where various drones are used in combat to conduct unmanned surveillance and attacks, and in recent years the US military has trained drones pilots [1]. In addition, it is used in various fields such as forest fire monitoring, border surveillance, broadcasting, aerial surveys, environmental monitoring, and courier delivery. In recent years, the range of applications for personal use and leisure has been widening. The next most frequently used field is the automotive sector. In the case of

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 364–369, 2020.
https://doi.org/10.1007/978-981-32-9244-4_52

vehicles, autonomous navigation is added to general automobiles, and unmanned vehicles are expected to be realized within 10 years. On the other hand, the use of drones in the marine sector is somewhat lower than in aviation or the field of vehicles. In particular, many companies are conducting research and development in the aviation and automobile fields, but only a few representative companies are conducting research and development in the marine field. Conventional marine drones are mostly utilized in the public and military fields, but MR-SENTINEL proposed in this paper is a system used for private sector such as aqua farming and marine environmental monitoring. The MR-SENTINEL designed and developed in this paper uses an electric propulsion motor with a length of 3 m. MR-SENTINEL's communication system basically uses WiFi-Mesh and enables free operation without additional communication fee. In addition, existing systems used expensive sensors, but MR-SENTINEL uses low-cost sensors to lower the price, so it can be widely used in the private sector.

2 Related Works

2.1 ROS

ROS is short for Robot Operating System and is open source software [2]. It has been continuously updated since its first release in 2008. ROS provides hardware abstraction, sub-device control, general functionality, and messaging between processes, package management, and various debugging tools needed to develop robotic applications. ROS is basically a software platform and supports heterogeneous hardware at the same time. ROS has supported various hardware platforms such as ARM, MIPS, and x86, and the supported development languages are C, C++, Python, and Java. Robots based on ROS include Turtlebot 2, Pepper, Gatper, and Open Manipulator. Nevertheless, there is very little use of ROS in the USV field.

2.2 Autonomous Control Level

Drones are developed for the purpose of being used from remote control to fully autonomous operation. In the case of remote control, only a communication system for controlling the drone at a long distance is basically needed. The status of the drone itself and the surrounding environment information are determined by a user who is located in remotely.

Sensors for collecting environmental information are installed in "Remote Control with Vehicle stat knowledge" is one step higher than "Remote Control". Sensors such as a GPS, a camera, LiDAR, sonar, etc. is installed to collect the current location of the drone and the surrounding environment information. Sensor systems are limited to providing current status information to the operator in real time. There are also eight more levels.

3 MR-SENTINEL

3.1 Hull Design and Engineering

The unmanned surface vehicle system designed and implemented in this paper is named MR-SENTINEL. MR means Marine Robot. MR-SENTINEL is designed for monitoring the aquafarm or near the coastline and aims to operate the unmanned surface vehicle within 1 $\sim$ 2 km maximum based on the remotely located device. For this purpose, the hull was designed as a catamaran type with a length of 3 m and a width of 1.5 m.

For the analysis of the kinematic performance, a model-2 ship with ship lines and body plan was used. The loading conditions were added to the 12th loading condition of the Design Load Condition. Numerical modeling is the data obtained by idealizing the fluid in contact with the hull as a 4-node or 3-node fluid element, and using this, the motion analysis is performed using WAMIT (Ver. 5.3S) programmed with BEM (Boundary Element Method), And the results of the analysis are plotted according to the period (s) of the waves. The motion of the ship was calculated by applying the 6 - DOF motion to 13 conditions of wave incidence in 13 directions and 39 wave periods.

3.2 Hardware Specification

As mentioned in the previous section, MR-SENTINEL used Catamaran. The reason is that the use of a considerably sized WiFi-Mesh (13 “X13” X1 “) antenna, which is capable of telecommunication over 10 km, as a communication protocol for remote control. It includes peripheral equipment parts such as gimbal for the stable communication environment even in swaying and pitching of the hull. The reason for using WiFi Mesh is that the main consumer, fishermen, do not like it because of monthly charges for LTE or satellite communications that can be used for remote control.

The main power unit of MR-SENTINEL uses electric motor for clean sea environment. Electric motor specification is DC 24 V, 30A and BLDC type motor is used. In order to secure a certain level of torque to cope with the propulsion and external impact, a speed reducer was installed and the maximum number of revolutions of the propeller was determined to be 600 rpm. Two cameras were installed for marine surveillance. One is used to detect obstacles in the front and to check the surrounding conditions, and the other is for detecting marine jellyfish and garbage. In order to recognize jellyfish and garbage, we used image recognition technique using YOLO technique. In order to prevent collision at close range, two USD-20LX-H01 LiDAR sensors of HOKUYO Automatic Co., Ltd were used to detect obstacles within 10 m.

There are two main reasons for using the gyro sensor in MR-SENTINEL. The first reason is to reduce the error of positional information when MR-SENTINEL is not moving on the surface. Another reason is that it is used to assist in the processing of images acquired through the camera when the MR-SENTINEL’s pitch and rolling occur due to waves. In the case of the images obtained in the rolling and pitching situations, the horizon is formed in a higher region, or when an object inclined or partially cut out on the screen occurs.

3.3 Design and Implementation of MR-SENTINEL

MR-SENTINEL requires a suitable operating system to operate. Although there are various open source and commercial platforms such as ROS [2], ArduPilot [3], OPENROV [4], WARTHOG [5], and DroneCode [6], These are not suited to marine environment, so we decided to develop an operating system for MR-SENTINEL. The MR-SENTINEL system implemented in this paper consists largely of MR_SENTINEL_S, MR_SENTINEL_L, and Framework parts. MR_SENTINEL_S is installed on the unmanned surface vehicle, MR_SENTINEL_L is installed on the remote control system on the ground, and Framework is the common library used by MR_SENTINEL_S and MR_SENTINEL_L. The development language was developed with C# and .NET Framework 4.5.

The framework package consists of extracting functions commonly used in MR_SENTINEL_S and MR_SENTINEL_L. A GPS package for processing GPS signals, a Model package for storing files in a database or file system, a serial package for reading data from joysticks and gyroscopes, and a basic protocol for exchanging data between MR_SENTINEL_L and MR_SENTINEL_S A protocol package and a Net package, a Helper for reading configuration or content from a text file or database, and a UI package for defining a user interface.

MR_SENTINEL_S is package installed on unmanned surface vehicle itself. It is composed of Collect, Config, Frm, Gpio, Model, Serial, Server, UI and Worker. MR_SENTINEL is composed of a Collect package that reads data from sensors, Form package and UI package that provides user interface screen, Model to save text file or database configuration file, Serial package for controlling electric motor through RS-232C or RS-485, a Gpio package that directly controls the digital input/output device, and a Server package that allows access from MR_SENTINEL_L.

The basic functions of MR_SENTINEL_S are as follows. First, analyze the command from MR_SENTINEL_L to turn the left/right electric motor left and right. Because it is a catamaran, the left and right motors rotate in opposite directions, but there is no hindrance to performing forward and backward functions depending on the shape of the propeller.

Second, data is read from the gyro sensor, LiDAR and GPS connected to MR_SENTINEL_S and transmitted to MR_SENTINEL_L.

Third, MR_SENTINEL_S can initialize the electric motors, change the default setting, and rotate the motor directly left/right.

MR_SENTINEL_L is installed in a PC or a notebook computer for remote control, and exchanges control command and status information with MR_SENTINEL_S using the WiFi-Mesh. The MR_SENTINEL_L package consists of Client, Common, Config, Frm, Resources, Serial, Server, Worker, and Model. Most packages are similar to MR_SENTINEL_S, and the Resources package is for controlling and gathering images from the cameras. MR_SENTINEL_L also has a gyro sensor because MR_SENTINEL_L does not move in most fixed positions. As you already know, the GPS sensor has the problem that if it does not move in a stationary state, the effective radius to indicate its position will continue to grow, losing its position later on. Also, MR_SENTINEL_L performs object recognition by directly calling image data from two cameras as RTSP. MR_SENTINEL_L It has self-navigation function of LEVEL 6

level and image recognition function of LEVEL 7 level. The autonomous navigation function allows the user to determine the current location, then integrate the GPS, gyro sensor, and sea chart information to the destination and set the route, and then control the motor to arrive. Figure 1 shows user interface of MR_SENTINEL and an USV which is autonomous sailing.

Fig. 1. A user interface of MR_SENTINEL and an USV which is autonomous sailing

4 Conclusion

In the case of aviation drone, DJI's various product launching strategies have reached a stage where each individual can utilize many. In the case of unmanned ground vehicles, it has reached a level that can be offered as an additional service to commercial vehicles. It means that the big market has been fully built. However, in the case of the USV as a maritime drone, most of them are still operated for military and academic purposes, while commercial ships have reached the experimental level using a prototype. In particular, next year, YARA Birkeland, the world's first 100-m electric propulsion ship, will begin experimenting in Norway. However, in the case of the private small boat sector, there is not enough market. In this paper, we designed and implemented MR-SENTINEL, an unmanned surface vehicle system that can be used by each individual. MR-SENTINEL is a prototype and will continue to develop technology to the extent that it can be used for general individuals or fishermen to monitor the aquafarm and marine environment compared to existing products through continuous experiment.

References

1. Boyle, M.J.: The costs and consequences of drone warfare. Int. Affairs **89**(1), 1–29 (2013)
2. Quigley, M., Gerkey, B., Conley, K., Faust, J., Foote, T., Leibs, J., Berger, E., Wheeler, R., Ng, A.: ROS: an open-source robot operating system. In: Proceedings of IEEE International Conference on Robotics and Automation, pp. 12—17. IEEE Press, Kyobe (2009)
3. Bin, H., Justice, A.: The design of an unmanned aerial vehicle based on the ArduPilot. Indian J. Sci. Technol. **2**(4), 12–15 (2009)
4. Berlian, M.H., Sahpurta, T.E.R., Ardi, B.J.W., Dzatmika, L.W., Besari, A.R.A., Sudibyo, R. W., Sukaridhoto, S.: Design and implementation of smart environment monitoring and analytics in real-time system framework based on internet of things and big data. In: Proceedings of International Electronics Symposium, Denpasar, Indonesia, pp. 29–30 (2016)
5. Lang, R.G., Aguire, H., Batista, M.R., Dos Santos, P.H.N., Afuso, S.K., Massuda, W.S., da Silva, M.O., da Silva, I.N., Romero, R.A.: Description of the Warthog Robotics 2012 Project. In: 2012 RoboCup (2012)
6. Casagrande, K., Friederichs, J., Gonzalez, C., Humphries, T., Tinall, Z.: Cyber attacks and defense framework for unmanned aerial systems (UAS) environment. J. Unmanned Aerial Syst. **3**(1), 37–58 (2017)

A Study of Trim Tab Management System for an USV and a Small Boat

NamHyun Yoo(✉)

Department of Computer Science and Engineering,
KyungNam University, 7 Kyungnamdaehak-ro, Masanhappo-gu,
Changwon-si, GyeongNam 51767, Republic of Korea
hyun43@kyungnam.ac.kr

Abstract. In order for the ship to operate safely on the water surface, proper engine and efficient linearity are required. The engine provides adequate power to allow the vessel to travel to the desired destination, and efficient linearity helps reduce fuel consumption and safe operation of the fluid. Because large ships are heavy and the length of the hull is long, even if they are operated at high speed in the early stage, the bow does not go up and the stern is not pitch down. On the contrary, the small ship has a shorter hull length than the weight of the hull, which causes the bow to go up and the stern to go down at the beginning of the operation. In this case, the biggest problem is that the bottom of a small boat and the water surface are contacted is widened, causing a problem of a lot of resistance. When the hull resistance becomes large, the fuel consumption becomes large. In addition, it causes problems such as greenhouse gas emissions. The equipment used at this time is the trim tab. The trim tab is the hardware for improving the posture control and navigation performance of the small boat, and is mostly operated manually. In this paper, we designed and implemented Smart Trim Tab Management System which can provide optimum posture control and navigation performance according to the hull posture by applying artificial intelligence to manually operated trim tab system.

Keywords: Trim tab · Interceptors · Pitching · Small boat · Smart control

1 Introduction

In order for the ship to operate safely on the water surface, proper engine and efficient linearity are required. The engine provides adequate power to allow the vessel to travel to the desired destination, and efficient linearity helps reduce fuel consumption and safe operation of the fluid. Because large ships are heavy and the length of the hull is long, even if they are operated at high speed in the early stage, the bow does not go up and the stern is not pitch down. On the contrary, the small ship has a shorter hull length than the weight of the hull, which causes the bow to go up and the stern to go down at the beginning of the operation. In this case, the biggest problem is that the bottom of a

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 370–376, 2020.
https://doi.org/10.1007/978-981-32-9244-4_53

small boat and the water surface are contacted is widened, causing a problem of a lot of resistance. When the hull resistance becomes large, the fuel consumption becomes large. In addition, it causes problems such as greenhouse gas emissions. The equipment used at this time is the trim tab [1–4]. The trim tab is the hardware for improving the posture control and navigation performance of the small boat and USVs (Unmanned Surface Vehicles), and is mostly operated manually.

In this paper, we designed and implemented Smart Trim Tab Management System which can provide optimum posture control and navigation performance according to the hull posture by applying artificial intelligence to manually operated trim tab system. When using this system, it is possible to reduce the turning radius during rotation, and to reduce the fuel amount by reducing the hull resistance occurring at the beginning step of the sailing.

2 Related Works

The theoretical background regarding the trim tab can be found in the paper [1]. The key factors related to the trim tab presented in [1] are the length and weight of the hull, the size and model of the transom, the size and angle of the trim tab, the total resistance and the speed of the ship. Trim tabs have some problems to be used in high-speed ships or large ships. When the hull is enlarged due to the plate type trim tab, it is difficult to construct a system for this purpose, and it is also ineffective from the mechanical point of view. A system designed to solve this problem is an interceptor. The interceptor system is more difficult to manufacture than the plate-based trim tab system, but has the advantage that more detailed control is possible [5–7]. However, the trim tab system is efficient for small ships. The system structure is simple, easy to manufacture, and easy to operate. Also, it has a merit that the price is low because of its simple structure.

3 Design and Implementation Trim Tab Management System

3.1 Equations for Trim Tab Management System

The basic equations for the trim tab management system implemented in this paper is based on [1] and [2]. Equation begins with determination of plate lift increment. This parameter is added to the total displacement of the planing hull which is used to calculate the displacement.

$$\Delta_F = 0.14025 L_F \delta \sigma b \left(\frac{\rho}{2} V^2\right) \tag{1}$$

Here, LF (m) is the flap chord, b (m) is the beam of planing craft, σ is the flap span-beam ratio, ρ (Kg/m^3) is the density of the water and V is the velocity of the craft (m/s). By having ΔF, it is easy to take into consideration the usage of trim tab for a planing

hull. This basically means that it is now possible to calculate the new displacement and new LCG (Longitudinal Enter of Gravity) of the planing body, as follows:

$$\Delta_{\varepsilon} = \Delta - \Delta_F \tag{2}$$

$$LCG_{\varepsilon} = \frac{(\Delta \times LCG - 0.6 \times \Delta_F \times b)}{\Delta_{\varepsilon}} \tag{3}$$

Where Δe and LCGe are called the Effective Displacement and Effective LCG, respectively. It is then possible to start the calculations of the original Savitsky's method [1] in which the mass and LCG in the governing equations should be replaced by the new Mass and LCG which considers the inclusion of trim tab in the planing body.

The original Savitsky's method [1] includes 30 steps which will be studied and executed. For getting the calculations started, first a trim angle must be guessed for the planing body, and for terminating the calculations, the suitability of this guess should be assessed.

Subsequently, it is necessary to find the involved constants. These constants include the beam Froude number, CV, and the lift coefficient CLβ. Lift coefficient is the parameter which gives enough force to lift the planing body. These two constants are calculated by Eqs. (4) and (5).

$$C_V = \frac{V}{\sqrt{gB}} \tag{4}$$

$$C_{L\beta} = \frac{mg}{0.5\rho V^2 B^2} \tag{5}$$

Here, g is the gravity acceleration, Cv is the speed coefficient, and m is the mass of the craft. By having the lift coefficient and utilizing Eqs. (6) and (7), it is possible to calculate the length-beam ratio, λ, as in

$$C_{L\beta} = C_{L0} - 0.0065\beta C_{L0} \tag{6}$$

$$C_{L0} = \tau^{1.1}\left(0.012\lambda^{0.5} + \frac{0.0055\lambda^{2.5}}{C_v^2}\right) \tag{7}$$

Where C_{L0} is the lift coefficient of planing plate, β is the deadrise angle of the craft (degrees), τ is the trim angle (degrees) of the planing surface, and λ is the non-dimensional mean wetted length of the planing boat. Next step is to find frictional Drag, D_f, but beforehand, a mean velocity (V_m) ought to be calculated for substituting in the equation of frictional drag. The mean velocity can be calculated from Eq. (8) as in.

$$V_m = V\left[1 - \frac{0.012\lambda^{0.5}\tau^{1.1} - 0.0065\beta\left(0.012\lambda^{0.5}\tau^{1.1}\right)^{0.6}}{\lambda\cos\tau}\right]^{0.5} \tag{8}$$

Friction drag coefficient (Cf) can be determined by ITTC-57 equation given in Eq. (9). After finding the Reynolds number, the frictional drag can be calculated. It is important to note that due to the ITTC standard, a quantity ΔCf = 0.0004 should be taken into consideration for calculating the frictional drag in Eq. (10).

$$C_f = \frac{0.075}{(\log_{10} Re - 2)^2}, \ \mathrm{Re} = V_m \lambda \frac{B}{\upsilon} \tag{9}$$

$$\mathrm{D}_f = 0.5 \frac{\rho V_m^2 \lambda B^2}{\cos \beta} \left(C_f + \Delta C_f\right) \tag{10}$$

Where Re is the Reynolds number and ν is the kinematic viscosity of the water. Later, forces that have to be calculated are hydrodynamic force normal to the bottom, i.e. N, and total drag, i.e. D, which are presented in Eq. (11) and (12).

$$mg = N \cos \tau \rightarrow \mathrm{N} = \frac{mg}{\cos \tau}, \tag{11}$$

$$\mathrm{D} = \mathrm{mg} \tan \tau + \frac{D_f}{\cos \tau} + 0.0052 \Delta_f \tan(\tau + \delta) \tag{12}$$

Here, the first term is induced drag of the craft, second term is frictional drag, and the last term is the drag of the flap. Moreover, the distance of frictional drag from VCG, i.e. a, and the distance of normal force from LCG, i.e. c, have to be determined. For this purpose, finding the center of pressure, Cp, is imperative. These parameters are presented in Eq. (13) through (14).

$$C_p = 0.75 - \frac{1}{\frac{5.21 C_V^2}{\lambda^2} + 2.39} \tag{13}$$

$$\mathrm{c} = \mathrm{LCG} - l_p = LCG - C_p \lambda \mathrm{b}, \ \mathrm{a} = \mathrm{VCG} - \frac{b}{4} \tan \beta \tag{14}$$

With all the stated forces calculated, it is now possible to calculate the pitch moment of the planing body from the following equation:

$$M_{tot} = mg \left[\frac{c}{\cos \tau} (1 - \sin \tau \sin(\tau + \varepsilon) - f \sin \tau)\right] + D_f (a - f) \tag{15}$$

Where ε is the inclination of thrust line relative to the keel (degree). Parameter f is the vertical distance of the thrust from center of gravity. For reaching the equilibrium, Eq. 15 should be set equal to zero. Normally, this cannot be done by guessing and it has to be calculated by an interpolation scheme. In Eq. 15, if Mtot is negative, it implies that the guessed trim angle at the first step is too low. Therefore, trim angle should be increased and all the calculations must be repeated again. This process continues until Mtot becomes a positive value.

3.2 Proper Tab Size for Trim Tab Management System

The plate that affect the posture of the hull by friction with the water surface is affected by the length and weight of the hull. For most boats, the width and weight are determined by the length of the hull, so the size of the existing trim tab depends on the length of the hull. Since the length of the hull tested in this paper is about 6 m, the plate size is determined as 9 "X9" suggested by [1] and existing commercial products, and an Ansys Fluent analysis was conducted to verify this suggestion.

In the case of an actuator that adjusts the water contact of the plate, an electric motor that can have a fast response speed according to the characteristics of a boat that needs to plane at high speed is used, and specifications are 12 v, 2.3 A, and 23,000 rpm. In addition, a ball screw type drive part is used for the up and down adjustment. When the motor rotates, the screw part is rotated through the reduction gear and the screw is moved forward and backward along the thread inside. An upper case for handling the cable from the motor and a lower case for connecting the lower screw housing were constructed and waterproofed using liquid gasket to seal the upper and lower connections.

The trim tab determines whether or not to operate depending on the posture of the hull. For this purpose, a sensor for measuring the posture of the ship is required. IMU is the most used sensor, but the price is so high that it is difficult to use in a small boat. In this paper, we used the tilt sensor and gyro sensor to measure the instantaneous posture of the hull. Especially, it is difficult to estimate the instantaneous hull posture because of the occurrence of pounding action during high speed operation. In order to solve this problem, the noise generated by the pounding is removed by the fusion of the tilt sensor data and the data of the gyro sensor, and the accuracy is improved. The specifications of the sensors used in this paper are: angular velocity ± 2000°/s, acceleration ± 16 g,

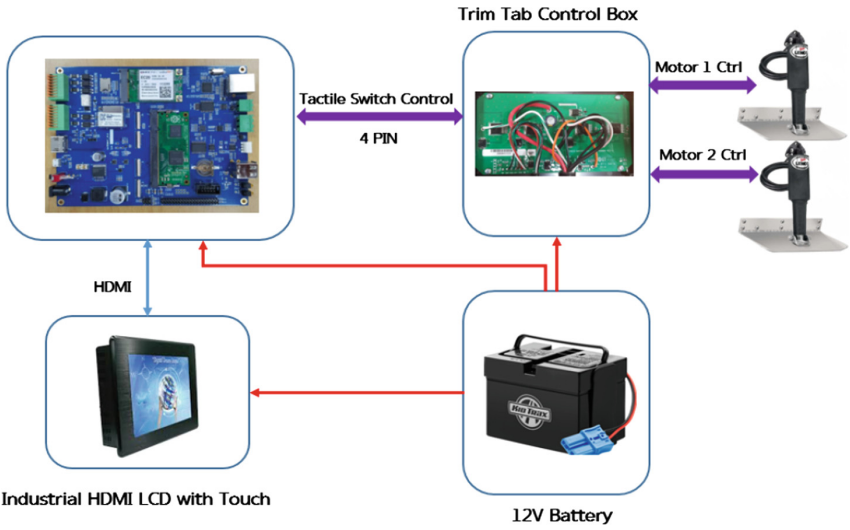


Fig. 1. The architecture of trim tab system.

self ±4800μT, temperature −10–80 °C, resolution 0.001° and Response Time <1 ms. The tilt and gyro sensor read the information once a second and uses the ARM 32 bit Cortex M series to provide the user interface to the user. The overall system configuration is shown in Fig. 1.

4 Analysis of Marine Experiments and Results

The effect of the trim tab is known to have the effect of increasing the safety, reducing the turning radius and reducing the fuel cost at high speed planning. In the case of the Smart Trim Tab Manager System, which is designed in this paper, it is necessary to verify the experiment by using artificial intelligence according to the posture of the ship rather than manually operating the boat operator. For this, six experiments were carried out from January to February 2018 at sea. The length of the boat used in the experiment was 6 m and the maximum speed was more than 20 knots. The marine wave height was 0.3 ∼ 0.5, and two tilt and gyro sensors were installed. DGPS was installed to measure the moving distance and speed. Experiments were carried out in the presence of external experts to determine whether the reduction of the turning radius and the reduction of the oil cost due to the decrease in the resistance at the early stage of the voyage were effective. As a result of averaging the results of 6 experiments, the radius of rotation decreased by 11.33% on the left side and by 12.82% on the right side.

5 Conclusion and Future Work

Most large vessels, except for warships, do not operate at high speed in the early stages of operation. Even at the stable stage after the initial stage, there is little rapid acceleration or emergency stop operation. On the other hand, small vessels or USVs frequently perform rapid acceleration and rapid landing. The trim tab system is effective for rapid acceleration or rotation of small ships.

The initial trim tab system was mostly manual based, but nowadays many automatic systems have been applied. A typical automatic trim tab system uses a tilt sensor to measure the posture of the hull. The automatic trim tab system using only the tilt sensor has a problem that the accuracy is lowered when the pounding action is occurring at high speed. In this paper, tilt sensor and gyro sensor are installed in a trim tab system suitable for small boats of 6 m or less. Using the tilt sensor and the gyro sensor at the same time can improve the noise reduction of the sensors and the accuracy at high speed. When the trim tab management system proposed in this paper is applied, the radius of rotation decreased by 11.33% on the left side and by 12.82% on the right side. Also, fuel costs can be saved by 5–10%.

References

1. Savitsky, D.: Hydrodynamic design of planing hulls. J. Mar. Technol. **1**(1), 71–95 (1964)
2. Ekman, F., Ribas, C.N., Rydelius, F.: Model for Predicting Resistance and Running Attitude of High-Speed Craft Equipped with Interceptors. Royal Institute of Technology, Stockholm (2016)
3. Seo, K.C., Gopakumar, N., Atlar, M.: Experimental investigation of dynamic trim control devices in fast speed vessel. J. Navig. Port Res. **37**(2), 137–142 (2013)
4. Castro-Feliciano, E.L.: Doctoral Thesis, Co-Design of Planing Craft and Active Control Systems. The University of Michigan, Ann Arbor, Michigan, pp. 27–38 (2016)
5. Ghadimi, P., Noni, A., Nowruzi, H., Dashtimanesh, A., Tavakoli, S.: Parametric study of the effects of trim tabs on running trim and resistance of planing hulls. J. Adv. Shipping Ocean Eng. **3**(1), 1–12 (2014)
6. Day, A.H., Cooper, C.: An experimental study of interceptors for drag reduction on high-performance sailing Yachts. J. Ocean Eng. **38**(8–9), 983–994 (2011)
7. Karimi, M.H., Seif, M.S., Abbaspoor, M.: An experimental study of interceptor's effectiveness on hydrodynamic performance of high-speed planing craft. J. Polish Marit. Res. **20**(2), 21–29 (2013)

Applying Transfer Learning into Recommendations in the Case of Data Deficient

Meiling Ge[1,2], Xiaoming Wang[1,2(✉)], Guangyao Pang[1,2], Xinyan Wang[1,2], Xueyang Qin[1,2], and Jiehang Xie[1,2]

[1] Key Laboratory of Modern Teaching Technology, Ministry of Education, Xi'an 710062, China
[2] School of Computer Science, Shaanxi Normal University, Xi'an 710119, China
{gemeiling, wangxm}@snnu.edu.cn

Abstract. Recommendation systems recommend things of interest to users by extracting hidden features in reviews. However, it is difficult for the recommendation system to make correct recommendations for those items with very few reviews. To solve this data deficient problem, we propose a Recommendation method based on Transfer Learning (RBTL). It is based on CNN model and it predict ratings by extracting features. We introduce the idea of transfer learning and minimize the distance between the source domain(with some reviews and ratings) and the target domain(with few reviews and ratings) by using the standard distribution distance metric named Maximum Mean Discrepancy (MMD) so that the model trained in the source domain can predict the rating in the target domain.

Keywords: Recommender systems · Rating prediction · Convolutional neural networks · Transfer Learning

1 Introduction

Nowadays, recommendation systems play an important role in our life. Many people use recommendation systems in the daily life such as shopping, reading news and watching movies. The traditional recommendation methods are mainly divided into three categories: collaborative filtering-based recommendation methods [1], content-based recommendation methods [2] and hybrid recommendation methods [3]. There exits some shortcomings such as sparse data, data deficient, and so on. In the recent years, the deep learning algorithms are applied to the recommendation field [4] to solve these problems. Zhu et al. [5] proposed a recommendation method based on aspect-aware latent factor model(ALFM) which learns users and items latent factors based on ratings and introduces a weighted matrix to associate those latent factors to alleviate the data sparse. GUAN et al. [6] proposed an Attentive Aspect-based Recommendation Model (AARM) which models the interactions between synonymous and similar aspects. They used a neural attention network which simultaneously considers user, product and aspect information

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 377–382, 2020.
https://doi.org/10.1007/978-981-32-9244-4_54

constructed to capture a user's attention towards aspects when examining different products. To tackle data sparse problem, Wu et al. [7] proposed content embedding regularized matrix factorization (CERMF) which adopts convolution neural networks to simultaneously generate the independent embedding representations for the users and the items to improve the accuracy of recommendations.

These methods can achieve better recommendation performance in the case of a large amount of data. However, the traditional recommendation system with data deficient cannot make accurate recommendations. Some review domains which are less frequently rated, lack sufficient data to build a reliable prediction model. Considering that, we propose a Recommendation method Based on Transfer Learning (RBTL) that predicting ratings can solve data deficient problem. In the process of training, we minimize the distance between the source domain and the target domain in order to predict the rating in the target domain. The main contributions of this work are summarized as follows:

(1) We introduce MMD to review text in order to solve the data deficient problem of recommendation. We minimize the distance between the source domain and the target domain so that the model trained in the source domain can predict the rating in the target domain.
(2) We use the idea of transfer learning in the recommendation system that training a model using the data of the source domain can predict the ratings in the other domain. That is, Transfer knowledge learned in one area to another.
(3) The experiment proves that we have improved a lot in using the Transfer Learning idea to solve the problem of data deficient in the recommendation system.

The remainder of this paper is organized as follows: The next section presents our problem formulation. Section 3 describes our framework. The experimental evaluation and result analysis are conducted in Sect. 4. Finally, Sect. 5 concludes this paper.

2 Problem Formulation

In this section, we illustrate how to solve the data deficient problem by our model. The recommendation system extract hidden features from reviews in order to recommend people their interested things. We try to use data from other domains to train a model that can improve the accuracy in our applications.

We use w represent user reviews, and y represent the truth rating $\hat{y}$ represent the predicted rating. $x_s \in X_S$ denotes data in source domain, $x_t \in X_T$ denotes data in target domain and $d(X_S, X_T)$ denotes the domain loss of the source domain and the target domain. Then the problem can be formalized as:

Input: w

Output: $\hat{y}$

$$st \begin{cases} \arg\min(d(X_S, X_T)) \\ \arg\min(y, \hat{y}) \end{cases}$$

In other words, we need to find a model that takes user review as input and outputs the rating. And we need to calculate the domain loss between the source domain and the target domain. At the same time, we should minimize the domain loss.

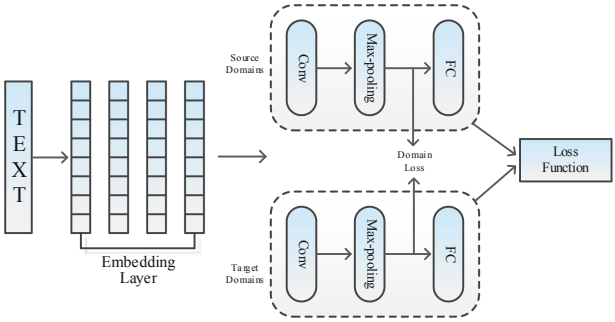


Fig. 1. The architecture of our model RBTL

3 Framework

In this section, we describe the details of our model. The architecture of the model is outlined in Fig. 1.

3.1 Basic Model

In the first layer, a word embedding function $f : M \rightarrow R^n$ maps the reviews of each word into a n dimensional vector, and then the reviews are represented as a matrix of word embedding with fixed length. The next layer follows the CNN model that contains a convolutional layer, a pooled payer, and a fully connected layer. In the first layer, we use filter $K_j \in R^{d \times t}$ on a window of words with size t to operator on word vectors V_j of user u. For V_j we perform a convolution operation regarding each kernel K_j by using ReLU [8] activation function.denoted as formula (1). The $*$ denotes the convolution operation and b_j denotes the bias. In the last layer, we use max-pooling to capture the most important feature-one with the highest value as the feature corresponding to this particular kernel denoted as formula (2). The final output of the convolutional layer is the concatenation of the output from its m neurons denoted as formula (3) and then been passed to a fully connected layer denoted as formula (4).

$$z_j = \mathrm{Re}LU(V_j * K_j + b_j) \tag{1}$$

$$o_j = \max\{z_1, z_2, \cdots z_{n-t+1}\} \tag{2}$$

$$O = \{o_1, o_2, \cdots o_m\} \tag{3}$$

$$X = WO + g \tag{4}$$

3.2 Minimize Domain Loss

To minimize the distance of source domain and target domain, we use MMD [9]. This distance is computed with respect to a particular representation $f(.)$. We define a representation $f(.)$, which operates on source data points, $x_s \in X_S$, and target data points, $x_t \in X_T$. Then an empirical approximation to this distance is computed as follows:

$$MMD(X_S, X_T) = \left\| \frac{1}{X_S} \sum_{x_s \in X_S} f(x_s) - \frac{1}{X_T} \sum_{x_t \in X_T} f(x_t) \right\| \tag{5}$$

Not only do we want to minimize the distance between domains, but we want to minimize the prediction error. To meet both these criteria is to minimize the loss:

$$L = L_C(y, \hat{y}) + \lambda MMD^2(X_S, X_T) \tag{6}$$

where $L_C(y, \hat{y})$ denotes the loss of the rating prediction, y and $\hat{y}$ denote the truth rating and the predicted rating, respectively. The hyper-parameter λ determines how strongly we would like to confuse the domains. To solve the parameters in the minimized loss function that we use Cross entropy loss function during the optimization learning process, we use Adaptive Moment Estimation (Adam). After training process, we can get the minimum loss L.

4 Experiment

4.1 Datasets

In our experiments, we use two datasets and they both from Amazon 5-core: Digital Music, and Musical Instruments. Digital Music is the larger dataset and it contains more than 16 thousand reviews, we take it as the source domain. Musical Instruments is the smaller dataset and it contains about 10 thousand reviews. We extract 2000 pieces of data as the target domain (to prove the target domain with few reviews and ratings). And the ratings of these datasets are integers in the range of [1, 5]. We randomly split the target domain dataset into training (80%) and test (20%) sets.

4.2 Experimental Settings

To evaluate the performance of rating prediction, we select a model that is not trained with the source domain but is directly trained and tested with the data in the target domain. We select model base_cnn as our baseline. In our proposed model RBTL, we use all the source domain data and the target domain data. In the process of training, we use Maximum Mean Discrepancy (MMD) to calculate the domain loss between the source domain and the target domain and minimize it. Then, we use the target domain to test our model. We reduce the data in target domain from 2000 to 1700, 1400, and 1000 for comparative experiments. Additionally, the learning rate is 2.5×10^{-5}. We set the epochs as 5 and the iterations is 50. We adopt the Accuracy to evaluate the performance of the algorithms.

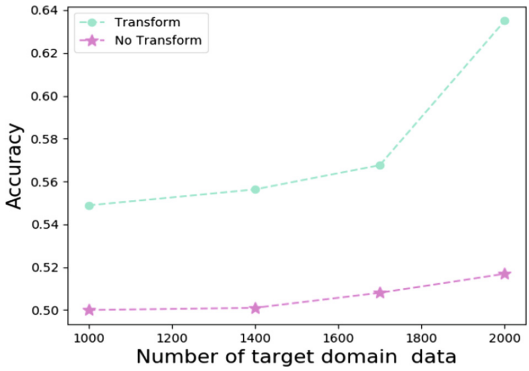


Fig. 2. The performance of our model RBTL

4.3 Performance Evaluation

The result is shown in Fig. 2. From the results, we find that there exists a relation between the accuracy and the number of target domain data. For the target domain data with a larger number of amount, the accuracy of the model is higher. And whatever the number of target domain data in the range of [1000, 2000], the accuracy of our model RBTL (transform) always higher than the model base_cnn (No Transform). It shows that our model gain about 12.05% improvement compared with the model base_cnn when the number of the target domain data is 2000.

5 Conclusion

It is shown that some review domains which are less frequently rated, lack sufficient data to build a reliable prediction model. In this paper, We introduce a transfer learning based model RBTL which can solve the problem of deficient data in Recommendation System. In the process of training we minimize the distance between the source domain and the target domain so that the model trained in the source domain can predict the rating in the target domain. And we use Maximum Mean Discrepancy(MMD) to calculate the domain loss between the source domain and the target domain. Experiment shows that our model performs better accuracy. In the future work, we will continue to improve our model accuracy of feature extraction by improving the neural network to better solve this problem.

Acknowledgments. This work was supported by the National Natural Science Foundation of China[Grant Nos. 61872228]; the Natural Science Basis Research Plan in Shaanxi Province of China [Grant No. 2017JM6060].

References

1. Collobert, R., Weston, J., Bottou, L., et al.: Natural language processing (almost) from scratch. J. Mach. Learn. Res. **12**(Aug), 2493–2537 (2011)
2. Wu, C.Y., Diao, Q., Qiu, M., et al.: Jointly modeling aspects, ratings and sentiments for movie recommendation. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD, vol. 14, pp. 193–202 (2014)
3. Nguyen, N.T., Rakowski, M., Rusin, M., et al.: Hybrid filtering methods applied in web-based movie recommendation system. In: International Conference on Knowledge-based & Intelligent Information & Engineering Systems. Springer, Heidelberg (2007)
4. He, X., Chua, T.S.C.: Neural factorization machines for sparse predictive analytics. In: Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 355–364. ACM (2017)
5. Cheng, Z., Ding, Y., Zhu, L., et al.: Aspect-aware latent factor model: rating prediction with ratings and reviews. In: Proceedings of the 2018 World Wide Web Conference on World Wide Web. International World Wide Web Conferences Steering Committee, pp. 639–648 (2018)
6. Wang, D., Xu, G., Deng, S.: Music recommendation via heterogeneous information graph embedding. In: International Joint Conference on Neural Networks. IEEE (2017)
7. Wu, H., Zhang, Z., Yue, K., et al.: Content embedding regularized matrix factorization for recommender systems. In: 2017 IEEE International Congress on Big Data, pp. 209–215. IEEE (2017)
8. Nair, V., Hinton, G.E.: Rectified linear units improve restricted Boltzmann machines. In: Proceedings of the 27th International Conference on Machine Learning, pp. 807–814 (2010)
9. Borgwardt, K.M., Gretton, A., Rasch, M.J., et al.: Integrating structured biological data by kernel maximum mean discrepancy. J. Bioinform. **22**(14), 49–57 (2006)

Opinion Similarity Regulated Public Opinion Network Embedding

Fei Ren, Xiaoliang Chen(✉), Yajun Du, Xianyong Li, and Ruomiao Li

School of Computer and Software Engineering, Xihua University, Chengdu, China
Fabulous_fly@163.com, xianyong@163.com, {chenxl,duyajun}@mail.xhu.edu.cn, Ithinkalso@126.com

Abstract. Developing a highly efficient transformer from an embedding public opinion network into a low-dimensional vector space contributes a lot to many research areas such as vertex classification, community detection and public opinion analysis, etc. Most existing network embedding methods have chosen to analysis in social networks. However, constructing a social network from public opinions is very sparsely, which would serve as an effective way to capture and process public opinions. On top of that, social network can only reflect the social relationships between nodes while the information derived from opinions is neglected. Hence, a network that incorporates opinion features of nodes into social networks is reported. This study evaluates the similarity of opinions from different nodes and connects them with enough similarity. The final public opinion network would certainly be denser than the social network. Experimental results show that researchers might give top priority to use the approach of public opinion network embedding compared with the regular social network methods, especially when the sentiment orientation of opinions is explicit.

Keywords: Network embedding · Opinion similarity · Public opinion network · Public opinion analysis

1 Introduction

Taking information from online social networks has become increasingly popular among users, especially college students, for their supplement of virtual and private space. Therefore, an increasing number of research [11, 15] have generally come to recognize the effectiveness of obtaining news and expressing opinions toward social events. These opinions involving different users will influence the trend of public opinion and how to predict this trend is a fundamental work of public opinion analysis.

This work is supported by the National Natural Science Foundation (Grant Nos. 61872298, 61532009, 61602389, 61802316, and 61472329), the Chunhui Plan Cooperation and Research Project, Ministry of Education of China (Nos. Z2015100 and Z2015109), the Civil Aviation Administration of China (No. PSDSA201802), the Chengdu Science and Technology Bureau (No. 2016-XT00-00015-GX), and Science and Technology Department of Sichuan Province (Nos. 2018026, 2017HH0083, 2016JY0244, and 2017RZ0009).

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 383–389, 2020.
https://doi.org/10.1007/978-981-32-9244-4_55

Network embedding plays an extremely important role in machine learning applications during the past decade. Network sparsity is the common problem appeared in network embedding tasks. Historically, research [1–4] investigating the problem associated with sparsity has focused on network structure. Early methods [1, 3, 6–10] include DeepWalk [1] that learns latent representations by using local information obtained from truncated random walks, in which the walks are treated as the equivalent of sentences.

First-order and second-order approximations are well considered from the study of method 'LINE' [3] that is reported by Tang et al. in 2015. Local and global network structures are preserved in the approach of LINE. In addition, researches [13, 14] have observed the global network structure. Only a matter of network structures is seen as a possible pattern to solve sparsity problems according to DeepWalk and LINE methods. Most work embraces the view that the two methods achieve some progress in network embedding while another information such as opinions from the tweet should not be neglected. Therefore, this study will focus on the opinion information from the tweet to deal with the sparsity problem. The attitudes that towards a real-world social event can easily obtained to construct a graph. However, the links between all the vertices are difficult to measure. This paper will measure the opinion similarity of all the vertices and detect the community by considering both the social relations and opinion information. All the links of a network depend on both opinion similarity and social relations, although the contribution of opinion similarity has differences in different kinds of networks. This work takes both the opinion similarity and social relationships into consideration simultaneously.

In summary, opinion features are introduced into traditional network embedding methods, and a public opinion network has proved helpful to solve the sparsity problem.

2 Problem Formulation

This section defines the problem of public opinion network embedding. The following three concepts are described.

Definition 1. Let $G_s = (V, E)$ denote a social network S, where $\mathrm{V} = \{u_n\}_{n=1}^{|N|}$ and $\mathrm{E} = \{e_{ij}|\mathrm{i} \in \mathrm{V}, \mathrm{j} \in \mathrm{V}\}$ represent the set of vertices and the links among these vertices, respectively. Each link e_{ij} can take one of the values between 1 and 0.

The vertices (rep. edges) of a graph are generally attached with a practical implication 'user' (rep. 'social links') to become a social network, in which the edge weights represent the binary status of the social link.

Definition 2. The similarity of opinions between two vertices is called the opinion similarity in social network. Let notations C_o and $\overrightarrow{o_i}$ denotes a corpus of opinions from all users in a social network and the vector form of user i's opinion, respectively. The opinion similarity between vertex i and j is defined as the Euclidean distance d_{ij} between $\overrightarrow{o_i}$ and $\overrightarrow{o_j}$.

An increasing number of tasks have benefited a lot from the vector form of a parameter in terms of simple matrix calculation. This work uses the approach word2vec [19] to train and calculate the vector form for any opinion.

Definition 3. A public opinion network is a three tuple $G_O = (V, E, S)$, where $\mathrm{V} = \{u_n\}_{n=1}^{|N|}$ represents the set of vertices of the network, $\mathrm{E} = \{e_{ij} | \mathrm{i} \in \mathrm{V}, \mathrm{j} \in \mathrm{V}\}$ reflects the social relationships between vertices i and j, and $\mathrm{S} = \{s_{ij} | \mathrm{i} \in \mathrm{V}, \mathrm{j} \in \mathrm{V}\}$ reflects the opinion similarity between vertices i and j.

3 Public Opinion Network Embedding

This section illustrates the method to establish a public opinion network. Furthermore, how to investigate the details and learn an expected low-dimensional representation with respect to an opinion network are presented.

Given a wide-spread public opinion event, we can easily find the users that comment on it and establish a social network $G_s = (V, E)$ by analyzing the comment relationship. Our goal is to find the mechanism of the mutual influence from comments to the event but not the interactions between vertices. Hence undirected graph is used to describe $G_s = (V, E)$.

On the other hand, a focused event will be set as a source node in the social network G_s. However, G_s. that is derived from a public opinion event represents the property of sparsity. A new concept called "Opinion Similarity" is introduced to solve this problem. People must hold a certain opinion towards the event. Hence, we can find the relationships between their opinions to complement their connections.

First of all, the opinions of the users that come from a network will be extracted as a corpus C_o. Then, the vector form of C_o should be prepared by using word2vec. We have the vector:

$$C_o \rightarrow s_i \in R^m, i \in (1, |V|), \tag{1}$$

where, s_i represents the vector form of user i's opinion. According to Definition 3, opinion similarity S_{ij} between user i and j can be denoted as:

$$S_{ij} = s_i - s_j, \tag{2}$$

On top of that, the parameter arithmetic average S_o of $S_{ij}(i, j \in V)$ should be calculated. In this study, S_o is defined as the threshold of the opinion similarity if $S_{ij} < S_o$ is satisfied. It is assumed that user i and user j have a high opinion similarity and they should connect to each other.

In this way, the opinion similarity of the whole network can be captured to establish a brand-new public opinion network $G_O = (V, E, S)$. The problem of public opinion network embedding is related to learning a function $f_O : V \rightarrow R^d$, in which the value of d far less than the modulus of the vertex set V, i.e. $\mathrm{d} \ll |\mathrm{V}|$. We define the joint probability p_1 between vector e_i and e_j by considering the fact that both social relationships E and opinion similarities S are preserved in the space R^d:

$$p_1 = \frac{1}{1 + exp(-e_i \cdot e_j)} \tag{3}$$

where e_i is the low-dimensional vector representation of vertex i in the social network. Correspondingly, the joint probability between opinion vector s_i and s_j can be represented as:

$$p_2 = \frac{1}{1 + exp(-s_i \cdot s_j)} \tag{4}$$

Then, we can obtain the following joint probability between vertex v_i and v_j:

$$p_0 = \alpha p_1 + \beta p_2 \tag{5}$$

where, parameters α and β represent the weight of social relations and opinion relations in a network, respectively. The empirical probability can be calculated according to the following equation $\hat{p}(i,j) = \frac{w_{ij}}{W}$ with $W = \sum_{(i,j)\in V} w_{ij}$. Finally, an expected objective function is presented as follow to preserve the social relations and opinion relations:

$$\mathrm{O} = \mathrm{d}(\hat{p}, p_0) \tag{6}$$

where $\mathrm{d}(\cdot,\cdot)$ denotes the distance between two distributions, which is usually called KL-divergence of two distributions. By finding the p_0 that minimizes the objective function in Eq. (6), the OSNE (opinion similarity network embedding) framework can represent every vertex in the network.

4 Experiments

4.1 Datasets

Three datasets are used to illustrate the effectiveness of this novel embedding algorithm.

(1) **SinaWeibo Hot Topic**: Making full use of Sinaweibo has become increasingly popular with people in China, where users can post tweets and retweet from others. Our group extracts the complete hot topic events during the whole 2018. The top five topics every month are chosen and prepared as a dataset with user information and opinion information.
(2) **Snapchat Friendship**: Snapchat is a photo-messaging application developed by camera company Snap Inc. Users can now enjoy the ease of making friends with each other and comment their photos with different attitudes in this plat. This study extracts all social relations from a relatively small and isolated group, which generates a network with approximate 11.5 thousand total users and 5 thousand edges.
(3) **Youtube network**: Youtube is a famous video website, which plays an important role in expressing public opinions when they watch any kind of videos in this

application. This research extracts the user information and user comments upside the video to construct a video-comment network.

Table 1 shows the detail nodes and edges with respect to the three datasets.

Table 1. Dataset statistic

Dataset	Nodes	Edges
Sinaweibo hot topic	**55.0 k**	20.0 k
Snapchat friendship	**11.5 k**	5.0 k
Youtube network	**35.5 k**	15.5 k

4.2 Quantitative Results

In the current study, opinion embedding will be demonstrated extremely useful among the accuracy of link predictions. In our experiments, 80% of the edges are randomly sampled as the training dataset. We compare with the following baseline results.

(1) **Matrix factorization for recommender systems** [12]. We apply their method to the social relation network to find the low-dimensional feature for users.
(2) **Large-scale information network embedding(LINE)** [3]. This method propose two types of node proximity, and we will compare with both of them.
(3) **Node2vec** [5]. This method learn the low-dimensional representations for vertices by the global network structure.

Our method outperforms all the baseline methods on link prediction tasks as is shown in Fig. 1.

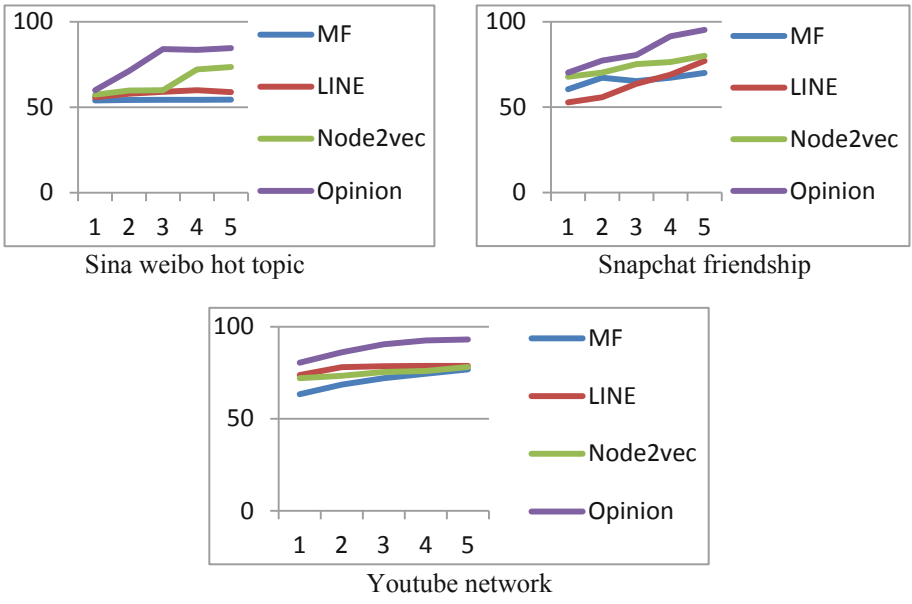


Fig. 1. Link prediction AUC (%) on all datasets

Table 2. Link prediction AUC (%) on all datasets

		Dimension of embedding K				
		2	4	8	16	32
Sinaweibo hot topic	MF	54.0	54.2	54.3	54.3	54.4
	LINE	55.8	57.9	59.0	60.0	58.8
	Node2vec	57.3	59.8	60.0	72.1	73.5
	Opinion	**60.0**	**71.2**	**84.0**	**83.5**	**84.6**
		Dimension of embedding K				
		2	4	8	16	32
Snapchat friendship	MF	60.5	67.2	65.3	67.2	70.0
	LINE	52.8	55.8	63.7	69.0	77.1
	Node2vec	67.8	70.2	75.2	76.4	80.1
	Opinion	**70.2**	**77.3**	**80.5**	**91.5**	**95.2**
		Dimension of embedding K				
		2	4	8	16	32
Youtube network	MF	63.4	68.6	72.1	74.5	76.9
	LINE	73.8	78.1	78.6	78.8	78.8
	Node2vec	72.1	73.4	75.5	76.0	78.2
	Opinion	**80.5**	**86.2**	**90.5**	**92.6**	**93.1**

5 Conclusion

This study was undertaken to design a novel public opinion network that is used to finding the low-dimensional features of opinion similarity. Then, this method has proved valuable to solve the sparsity problem in normal networks. Finally, the network structure and opinion similarity are combined to process network embedding.

The experiments results demonstrate that the opinion similarity increases the accuracy of link prediction in network embedding and makes up the features of social network. A natural progression of this work is to analyse the groups with highly opinion similarity. The issue of finding the group with highly opinion similarity is an intriguing one which could be usefully explored in further research.

References

1. Perozzi, B., Al-Rfou, R., Skiena, S.: DeepWalk: online learning of social representations (2014)
2. Wang, H., Zhang, F., Hou, M., et al.: SHINE: signed heterogeneous information network embedding for sentiment link prediction (2017)
3. Tang, J., Qu, M., Wang, M., et al.: LINE: large-scale information network embedding. In: 24th International Conference on World Wide Web, WWW 2015 (2015)
4. Cai, H.Y., Zheng, V.W., Chang, K.: A comprehensive survey of graph embedding: problems, techniques and applications. IEEE Trans. Knowl. Data Eng. 1 (2018)
5. Grover, A., Leskovec, J.: node2vec: scalable feature learning for networks (2016)

6. Chang, S., Han, W., Tang J, et al.: Heterogeneous network embedding via deep architectures (2015)
7. Yang, C., Liu, Z., Zhao, D., et al.: Network representation learning with rich text information. In: International Conference on Artificial Intelligence. AAAI Press (2015)
8. Chen, H., Perozzi, B., Al-Rfou, R., et al.: A tutorial on network embeddings (2018)
9. Ahmed, A., Shervashidze, N., Narayanamurthy, S., et al.: Distributed large-scale natural graph factorization. In: International Conference on World Wide Web. ACM (2013)
10. Belkin, M.: Laplacian eigenmaps and spectral techniques for embedding and clustering. In: Advances in Neural Information Processing Systems (2002)
11. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature (1998)
12. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. Computer **42**(8), 30–37 (2009)
13. Tenenbaum, J.B.: A global geometric framework for nonlinear dimensionality reduction. Science **290**(5500), 2319–2323 (2000)
14. Laurens, V.D.M.: Accelerating t-SNE using tree-based algorithms. J. Mach. Learn. Res. **15** (1), 3221–3245 (2014)
15. Abu-El-Haija, S., Perozzi, B., Al-Rfou, R., et al.: Watch your step: learning graph embeddings through attention (2017)

Short Text Sentiment Analysis of Micro-blog Based on BERT

Jianzhong Zheng, Xiaoliang Chen(✉), Yajun Du, Xianyong Li, and Jiabo Zhang

School of Computer and Software Engineering,
Xihua University, Chengdu, China
socool_zhong@163.com,
{chenxl,duyajun}@mail.xhu.edu.cn, xian-yong@163.com,
coolcool_bo@163.com

Abstract. Micro-blog has become increasingly popular among the general public. It has brought a lot of comment text to researchers for its great convenience, timely updating, and a wide variety of self-focused topics. Identifying the emotions expressed in these comments has become a valuable topic in order to make inferences for focused contents in Micro-blog. In this paper, we report on the effectiveness of the language representation model BERT [1] with respect to the sentiment classification tasks. Experimental results show that the pre-training of deep bidirectional transformers can improve the accuracy, recall and F1 score on sentiment classification. The final evaluation index of this problem by using a Github data set increased by 2.3% on average.

Keywords: BERT · Micro-blog · Sentiment classification · Short text · Pre-training

1 Introduction

Sentiment analysis is the process of evaluating texts with emotional color, which is a core research field in NLP. It is a meaningful process if we combine sentiment analysis with a hot social media: Microblog since identifying sentiment is considered to be the underlying technology for such specific applications as online portraits and public opinion analysis.

People like to submit comments and interact with others on Micro-blog applications such as Twitter and Sina. Therefore, a large number of short comments are produced, which are sparse but rich in emotional vocabulary compared to traditional official text. However, some difficulties appear in the analysis of comment sentiments due to their

This work is supported by the National Natural Science Foundation (Grant Nos. 61872298, 61532009, 61602389, 61802316, and 61472329), the Chunhui Plan Cooperation and Research Project, Ministry of Education of China (Nos. Z2015100 and Z2015109), the Civil Aviation Administration of China (No. PSDSA201802), the Chengdu Science and Technology Bureau (No. 2016-XT00-00015-GX), and Science and Technology Department of Sichuan Province (Nos. 2018026, 2017HH0083, 2016JY0244, and 2017RZ0009).

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 390–396, 2020.
https://doi.org/10.1007/978-981-32-9244-4_56

irregular sentence structures and limited length. The semantics of the word expression may differ from their usual meanings. Researchers prefer to enjoy the advantages of analyzing these comments since the sentiment implied in them will be a significant supplement to understanding the content and opinion of a micro-blog.

There are two main research ideas for sentiment classification: sentiment lexicon [2] and machine learning [3]. They use the sentiment lexicon-related evaluation or learn the combined features of text emotion to obtain the emotional polarity of the text, respectively. An increasing number of researchers are faced with a dilemma: should they require their work to concentrate on construction sentiment lexicon by traditional methods or should they make efforts to compute the representative features of text emotion by machine learning. This has aroused a lot of concern and heated debate. Machine learning method has been well accepted and acclaimed for their obvious advantages. Sentiment can be well classified by learning the implicit emotional features that are difficult to express in the dictionary. A typical machine learning model, recursive autoencoders network (RAE) (Socher et al.) [4] is reported to transfer each sentence to a low-latitude vector representation that can be used to classify text emotions and obtain an acceptable score. Researchers have been making every possible effort to update the existent score in this work. Therefore, two typical deep learning models, recurrent neural network (RNN) [5] and convolutional neural network (CNN) [6] were used for sentiment analysis and finally achieved a large performance boost. Compared with the RAE model, RNN and CNN can better understand the semantics and hence have better results for the emotional classification task.

Some improved models of deep learning have achieved better results. Tag-guided recursive neural network (TG-RNN) and tag-embedded recursive neural network (TE-RNN) [7] have advantages in learning and combining label vector and word vector. Most important, the subsequent model: LSTM [8] achieved a significant performance improvement again. LSTM continuously interacted before and after words in the compound operation process. A lot of work was adjusted within the framework of this mechanism to achieve a slight performance improvement [9]. Since then, there has failed to leaping work. A new language representation model that is called BERT [1] was proposed by the Google AI language team in 2018. This model is reported to capture the general performance gains with respect to some research issues in the field of natural language processing, which does not report the application evaluation result for short text sentiment analysis problems. Hence, our work attempts to use the BERT model for sentiment analysis.

2 Related Work

In the field of image processing, the pre-training of deep learning can significantly promote the effect of related image applications. However, deep learning directly applied to early NLP work was not satisfactory due to lack pre-training techniques.

Word embedding values are mainly captured by the neural network language model (NNLM) [10] and Word2vec [11]. NNLM constructs an n-gram language mode by design a three-layer neural network, which is used to predict the words that appear in a sentence, and its intermediate output of training is the word vector. On the other hand,

Word2vec uses CBOW and skip-gram to train the model to get the target word vector. Word embedding can be used to calculate inter-word relationships and are usually designed as the hidden layers of neural networks.

Word embedding is usually used at the bottom of a training process in an NLP task, in which 'Frozen' and 'Fine-Tuning' are two types of training. The former keeps the parameters that are loaded by the underlying network unchanged during the training process, and the latter changes with the appearance of updated training set.

However, word embedding does not correctly make a distinction between two polysemous words. The defect limits the improvement of the training effect. Thereafter, a new feature-based model ELMO [12] is recorded, which updates word embedding through a two-stage training method. This model contains two LTSM layers that compute the word embedding values with respect to syntactic and semantic information, respectively. This work first contributes a method to produce word vectors that can distinguish synonym one.

Google proposed a model framework 'transformer' [13] in 2017, which introduce pre-training on feature extraction for machine translation tasks. This work extended the early self-attention model to a multi-head structure. This new language model updates the best score for multiple NLP tasks, which training process is considered to be a typical case of Pre-Training.

Transformer has advantages over traditional LSTM in feature extraction and feature fusion. Pre-training model GPT (generative Pre-Training) [14] was proposed in 2018. This work is intended to construct a general learning approach to multi-domain NLP tasks by combining unsupervised 'pre-training' and supervised 'fine-tuning'. However, its one-way learning process is not conducive to integrate with the content features below the attention point, which limit its practical application scenarios.

In 2018, department Google AI Language adopts a two-stage training approach that is similar to GPT model. A bidirectional language model that is related to ELMO is applied in their pre-training phase. Finally, a model called BERT [1] is reported for pre-training of deep bidirectional transformers. BERT has strong universality and can be fine-tuned to most tasks of NLP such as sequence labeling, text classification, and so on.

The idea of Bert is to randomly remove the words on an input, and then the algorithm learns the language model for predicting missing words according to context-before and context-after of the input. Bert is mainly a typical masked language model (MLM).

The nice thing about this is that the representation that you've learned is that you can fuse contexts in both directions. Sentiment analysis task is to analyze a single sentence, one of the required information is to extract the emotion word feature vector in the sentence. Since each token has context representation in MLM pre-training, BERT can help us better extract emotional word features. In the affective analysis, as the characteristic word of affective words, the proportion of adjectives to the total sequence should be calculated first. Secondly, negative words need special treatment when extracting features. Because negative words are of vital importance and directly affect the emotional tendency of a sentence, prefix NOT_ is added to all words between negative words and the nearest punctuation, which can give full play to the role of negative words.

Its input part is a linear sequence, the sentence sequence is separated by a separator, and two identifiers are added at the beginning and the end. Each word has three embedding: embedded location information, this is because the word order is very important in the NLP characteristics, need to be here to encode the location information; The word e embeds, this is what we've been talking about embeds; The third is sentence embedding, Because in our emotions are all clauses in the analysis task input, so the sentence as A whole embedding item corresponding to every single word, A embedded here we only use the sentence, and then put the word A superposition of three embedding, Bert input can be built up. In order to obtain the emotion classification result, we represent the input sequence, and use the final hidden vector C corresponding to the first input tag [CLS] as the aggregation representation.

3 Experiment

This paper analyzes the effect of Bert applied for sentiment analysis by the following experiments. The data set is from https://github.com/1007105767/WeiboData, in which contains 180,000 training sets and 15,000 test sets. The data comes from Weibo's commentary short text, and the emotional word data is shown in Table 1.

Table 1. Experimental data information

Sentiment Type	dismay	disgust	anger	joy
Number of data	26000+	42000+	49000+	57000+

This experiment makes the following steps of in order to adopt Bert to the sentiment classification of in Micro-blog text:

(1) Load the Chinese pre-training model of Bert by the code: fine-tuning, which is given by Google in [1]. The experiment handles the sentiment classification task by calling the Python file 'run_classifier.Py'. We adjust the relevant parameters about 'read file' by calling the base class 'DataProcessor'. Considering the difference between the format of experimental data and the format of model input, as well as the dichotomy principle of the sentiment classification problem, this paper re-defines the 0, 1, 2, 3 label method of the input file. The new file processing class is called 'WeiboProcessor'.

(2) Use the static method '_read_txt' to obtain the training set, test set and label of the experimental data in this experiment, where the training data and the test set data information have been given above. The test set data volume is 12 times that of the training set. The training set is divided into four separate files with labels. The specific label information is given in Table 2. Pay attention to the emotional tag customization when reading data.

(3) The model Bert needs to define four parameters: guid, text_a, text_b, and label, which mark the user ID, the training sentence set a, the training sentence set b, and the label set, respectively. The experiment retains three items except text_b (usually used for answering data of QA questions) by considering the characteristics of the sentence sentiment classification problem. The remaining model parameters are shown in Table 3.

The experimental results are shown in Fig. 1. It can be seen from the figure that the simple use of CNN for the emotional classification of microblog short text is the worst because its own convolution property is difficult for short text to capture its context features. For both GRU and LSTM, these two models are improved by the cyclic neural network RNN, and their effects should be similar. However, because the LSTM network itself is more complex in processing, its experimental effects are slightly better than its brother GRU network. For text, one-way information extraction seems to be insufficient, while two-way network is always better than one-way network in feature extraction, so bi-lstm always performs better. A CNN based convolutional memory network combined the advantages of CNN and memory network, has a good performance in the experiment, in the pure CNN network experimental results will improve the accuracy of 5.5% points. CNN was used for word convolution, and bi-lstm was used to extract features for training. The current best score was 88.92%.

Table 2. Training set data emotional information

File name	Sentiment type
Train 0	joy
Train 1	anger
Train 2	disgust
Train 3	dismay

Table 3. Model parameters of Bert on sentiment classification

do_train	true
do_eval	true
dopredict	true
max_seq_length	128
train_batch_size	32
learning_rate	5e-5
num_train_epochs	2.0

In this paper, the BERT model was used. The open source pre-training model was directly used for training, and then the test experiment was carried out. Finally, the accuracy of the results obtained was 91.23%. For weibo essay, the difficulty lies in how to improve the understanding of the sentence and emotional words and the context of the word, just the BERT benefited from the training of the large amounts of data and the appropriate frame structure, whether it is a long text and this essay can do better as a result, from the point of the experimental results, this paper USES BERT model to simulate the weibo this essay sentiment analysis, classification effect have got improved. Conclusion the available scope of reference 1 has been extended.

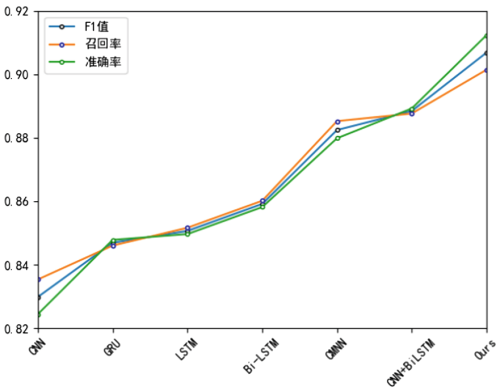


Fig. 1. Test comparison results

4 Conclusion

In this paper, we proposed the use of BERT to do emotional analysis of short weibo texts. The model can extract semantic features efficiently. In the pre-training process, the model architecture can fully integrate the semantic features of sentences and convert them into feature vectors with context information by word embedding, sentence embedding and position embedding. In the fine-tuning process, we adjust it to the training only for emotion classification, extract its emotional features, and then carry out the emotion classification of the short text. Experimental results show that BERT can be used to process emotion analysis tasks and get performance improvement. On the other hand, only one data set is used to test all the methods. Our further work will obtain multiple common weibo data sets on different topics for further testing, in order to seek the average improvement of BERT's emotional analysis task.

References

1. Devlin, J., et al.: Bert: pre-training of deep bidirectional transformers for language understanding (2018). arXiv preprint: arXiv:1810.04805
2. Hatzivassiloglou, V., McKeown, K.R.: Predicting the semantic orientation of adjectives. In: Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics. Association for Computational Linguistics (1997)
3. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up?: sentiment classification using machine learning techniques. In: Proceedings of the ACL 2002 Conference on Empirical Methods in Natural Language Processing, vol. 10. Association for Computational Linguistics (2002)
4. Socher, R., et al.: Semi-supervised recursive autoencoders for predicting sentiment distributions. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics (2011)
5. Elman, J.L.: Finding structure in time. Cogn. Sci. **14**(2), 179–211 (1990)

6. LeCun, Y., et al.: Backpropagation applied to handwritten zip code recognition. Neural Comput. **1**(4), 541–551 (1989)
7. Qian, Q., et al.: Learning tag embeddings and tag-specific composition functions in recursive neural network. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Long Papers, vol. 1 (2015)
8. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997)
9. Hinton, G.E.: Learning distributed representations of concepts. In: Proceedings of the Eighth Annual Conference of the Cognitive Science Society, vol. 1 (1986)
10. Bengio, Y., et al.: A neural probabilistic language model. J. Mach. Learn. Res. **3**, 1137–1155 (2003)
11. Mikolov, T., et al.: Distributed representations of words and phrases and their compositionality. In: Advances in Neural Information Processing Systems (2013)
12. Peters, M.E., et al.: Deep contextualized word representations (2018). arXiv preprint: arXiv:1802.05365
13. Vaswani, A., et al.: Attention is all you need. In: Advances in Neural Information Processing Systems (2017)
14. Radford, A., et al.: Improving language understanding by generative pre-training (2018). https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/language-unsupervised/language_understanding_paper.pdf

Situation Assessment with Random Bayesian Network Forest

Xiushe Zhang[1], Ming Guan[2], Xiaoquan Hu[1], Chunlei Han[1], and Jianshe Wu[1,2(✉)]

[1] The 20th Research Institute of China Electronics Technology Group Corporation, Xi'an 710068, China
jshwu@mail.xidian.edu.ch
[2] The Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, Xidian University, Xi'an 710071, China

Abstract. This paper presents a framework based on the random Bayesian networks forest (RBNF) to assess situation, which consists of three main components: data processing, offline construction, and training of the Bayesian network. Single Bayesian network is constructed by combination of the K2 algorithm and genetic algorithm. In addition, in order to improve the ability of generalization and increase prediction accuracy in universal situations, bagging learning is used to eliminate redundant information and extract implicit dependences between features captured by various sensors. The proposed framework is evaluated and tested in a real Australia credit data. Simulation results show that the framework can provide sufficient performance for situation assessment in terms of accuracy, reliability and robustness.

Keywords: Situation assessment · Bagging learning · Bayesian network · K2 · Genetic algorithm

1 Introduction

With the development of artificial intelligence, situation assessment (SA) is playing a more and more import role in our life and has many useful applications. SA can provide supportive advices to decision makers.

In the early stages of development, SA methods are mainly based on the template matching [1], static Bayesian network [2], D-S Evidence theory [3] and so on. In recent years, deep learning models (e.g., Convolutional Neural Network, Deep Belief Network and Recurrent Neural Network [4, 5]) have been applied in this field. The merit of deep learning is that it can mine the invisible relationships between features with nonlinear expressions. As a black box on the other hand, we cannot find the internal causal relationships from deep learning, but can from the Bayesian network. In addition, deep learning requires a large number of tagged data sets, but in reality it is difficult to obtain.

The advantage of the Bayesian network is that it clearly expresses the causal relationship of reasoning. The shortcoming is that it cannot find the invisible relationships between features and relies on expert knowledge. Refs. [6, 7] provide evolutionary

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 397–402, 2020.
https://doi.org/10.1007/978-981-32-9244-4_57

models of Bayesian networks, but none of them overcomes these problems. The main motivation of this paper is to overcome these problems to improve the accuracy of Bayesian networks.

The work of this paper has three aspects: (a) In order to eliminate the uncertainty brought by sensors, Fuzzy theory is introduced; (b) The K2 algorithm is combined with the genetic algorithm to learn the optimal structure of Bayesian network without expert knowledge. (c) Bagging learning is used to mine the invisible relationships between features and improve the generalization ability of the model.

This paper is organized as follows. Section 2 is the preliminaries. Section 3 describes the details of the proposed RBNF algorithm for SA. Simulation results are shown in Sect. 4. Section 5 concludes this paper.

2 Preliminaries

2.1 Bayesian Network

A Bayesian network is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a direct acyclic graph (DAG).

A Bayesian network is expressed as the form $G = (S, P)$, where S represents the structure of the Bayesian network, each node in S is a random variable, a state of a node represents a value of a random variable. Direct edges represent the dependencies between each pair of nodes. P is a conditional probability table used to indicate the relationships between node X_i and its parents $Pa(X_i)$. According to the conditional independences, the joint probability of the N variables in Bayesian network can be expressed by $P(X_1, X_2, \ldots X_N) = \prod_1^N P(X_i/Pa(X_i))$.

2.2 K2 Algorithm

The K2 is a classic algorithm for the structure learning of a Bayesian network, which is a greedy algorithm based on the hill climbing proposed by Cooper and Herskovits [8].

At initialization, the K2 algorithm defines an order of the nodes and specifies that each node's parents must be in front of it. Then in each iteration, the K2 algorithm selects the candidate parents for each node in turn. For a node, if a candidate parent node is added to the network, the Bayesian information criterion (BIC) score will be recalculated. If the new BIC score value is higher than the old one, the candidate node is added to the network. Finally, the network with the highest BIC score is selected as the final network.

The final network structure is determined by the order of the nodes. In this paper, we use genetic algorithm to find the optimal order of nodes.

3 RBNF Algorithm

In this section, we introduce the RBNF algorithm in details.

3.1 Data Processing

In order to eliminate the uncertainty brought by sensors, we do not give a confirmed value to every feature. According to the fuzzy theory [9], Each random variable R in the Bayesian network is associated with a set of possible values, i.e., $r = <s_1^{p_1}, s_2^{p_2}, \ldots, s_M^{p_M}>$, along with a possibility distribution P that reflects the uncertainty on the state of the random variables, where P_i is the associated probabilities such that $\sum_i P_i = 1$. P_i can be determined by formulation:

$$P(r = s_i) = \frac{f(r = s_i)}{\sum_{k=1}^{M} f(r = s_k)}, \tag{1}$$

where $f(r = s_i)$ reflects the determined probability of state s_i. Each state s_i of a random variable has a standard security value x_i', then the $f(r = s_i)$ can be shown as:

$$f(r = s_i) = \exp(-\frac{(x - x_i')^2}{2\delta_i^2}), \tag{2}$$

where δ_i is the confidence level (we usually set the value between 0.8 and 1), and x is the continuous measurement value collected by the sensor.

Due to the redefinition of each random variable, the respective fuzzy joint probability of Bayesian network $P(r, Pa_1(r), \cdots, Pa_k(r))$ can be simply calculated using a combination of the chain rule.

$$\begin{aligned} P(r, \mathrm{Pa}_1(r), \ldots, Pa_k(r)) = \sum_{i=1}^{N} \sum_{\mathrm{j}_1=1}^{M_1} \cdots \sum_{j_K=1}^{M_k} P(r = s_i | Pa_1 = t_{j_1}, \ldots, Pa_{_k} = t_{j_K}) \\ \times P(Pa_1 = t_{j_1}) \times \ldots \times P(Pa_k = t_{j_K}) \end{aligned} \tag{3}$$

in which $P(\mathrm{Pa}(r))$ is the state estimation of the parent, $P(r|Pa(r))$ is the conditional probability table (CPT), and M_1,..., M_k are the number of possible values of $Pa_1(r)$ to $Pa_k(r)$.

3.2 Construction of Bayesian Network

We combine the K2 algorithm with the genetic algorithm to find the best network as follows.

Initial population: Each individual in the population is an N-dimensional vector that represents the ordering of the nodes without repeat. L individuals are randomly generated. When initializing a population, we can manually add some good individuals to speed up the convergence of the algorithm. Calculate the Pearson coefficient between each pair of nodes. If the value of the Pearson coefficient is higher than the given threshold, which indicates that the correlation between the two nodes is relatively high, the two nodes may be the parent-child relationship. Since the parent of a node is in front of it, the more nodes a node has associated with, the higher the order of the node.

Crossover: The single point cross method is used to generate a new individual. Randomly select an individual from the group, then randomly select an intersection and exchange the left and right parts.

Mutation: Step1: set a small mutation probability value P_m and then generate a random decimal V. If $V < P_{m,}$ go to Step 2; Otherwise, nothing need to be done.

Step 2: when $V < P_{m,}$ randomly generate two mutation positions and then exchange values at these two positions.

Selection. The K2 algorithm determines the network structure according to the order of nodes; we use the Expectation-Maximization (EM) algorithm to learn the parameters (conditional probability) and the junction tree inference algorithm to calculate its prediction accuracy on the test set. Individuals with the top L accuracy from the population are selected for the next iteration.

After multiple iterations of the above process, the network with the highest accuracy in the population is selected.

3.3 Bagging Learning

Bagging learning on Bayesian network mines hidden relationships between features and improves the generalization capability.

The most important operation of bagging learning is to extract subdata sets from the sample dimension and feature dimension of the data set as training datasets. One subset can be used to construct a Bayesian network. By extracting subsets of data, some potential dependency can be discovered. The detailed steps are shown as follows.

Step 1: Setting the initial parameters, including the number of iterations T, the set of Bayesian networks $G = \phi$, the proportion of extraction P_e.
Step 2: Extract a subdata set from a given dataset by scale value P_e. The subdata set will be the training set.
Step 3: Utilizing the algorithm in Sect. 3.2 to generate a Bayesian network which has the best performance and adding it to the Bayesian network forest G.
Step 4: If the current number of iterations reaches the given number of iterations, the algorithm terminates; otherwise, repeat steps 2 and 3.

There will be T different Bayesian networks after T iterations, these T networks make up a forest. When test data are predicted through this forest model, there will be T results. The situation estimate is obtained by weighted summation of these T results.

4 Experimental Results

The real Australia credit data (http://archive.ics.uci.edu/ml/datasets.html) is used in the experiment. Each sample in the data has 15 attribute variables: 6 variables are continuous values, 8 variables are discrete values, and 1 variable is the label.

The proposed framework is compared with Bayesian network (BN) [2], Multi-Entity Bayesian network (MEBN) [7], Naive Bayes classier (NBC) [10], Artificial Neural network (ANN) [11] and Support Vector Machine (SVM) [12]. Accuracy (ACC),

False Negative rate (FN) and False Positive rate (FP) are chosen as the metric. This paper use 5-fold cross validation to evaluate the performance of these involved models. The compared results are in Fig. 1.

The experimental results in Fig. 1 show that the ACC value of the proposed algorithm is higher than other algorithms, while the FN and the FP values are lower than other algorithms. Since the higher the ACC value, the lower the FN and FP values, the better the performance of the algorithm, the algorithm proposed in this paper is superior to other algorithms.

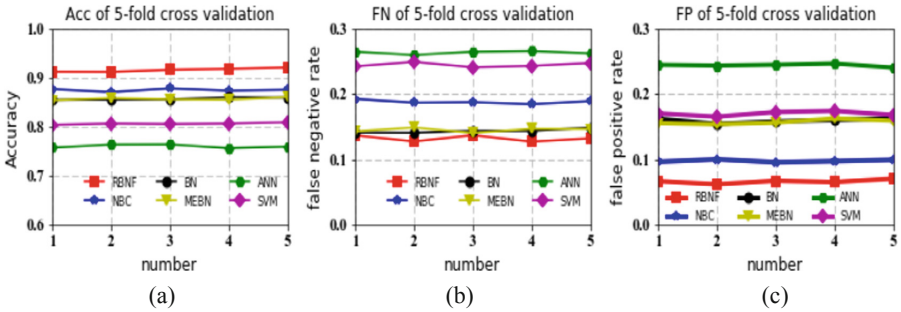


Fig. 1. (a) Accuracy of 5-fold cross validation; (b) False Negative rate of 5-fold cross validation; (c) False Positive rate of 5-fold cross validation.

The real-time prediction results of SA are shown in the Fig. 2(a). If the threshold is set to 0.5, the real-time prediction results are consistent with the label. Experimental results prove that the algorithm in this paper is effective. In terms of prediction time complexity, although RBNF algorithm integrates multiple Bayesian networks, it can be processed in parallel. On the contrary, because of data down-sampling, the network is simpler and faster than other Bayesian networks. The result is shown in Fig. 2(b).

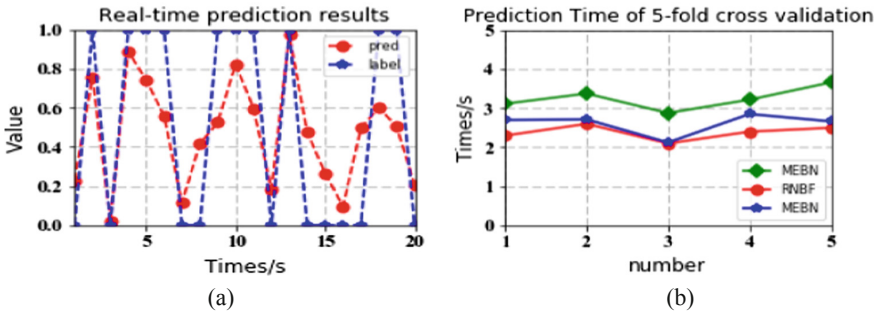


Fig. 2. (a) Real-time prediction; (b) Prediction time of 5-fold cross validation.

5 Conclusion

We propose a novel random Bayesian network forest (RBNF) algorithm to deal with situation assessment (SA). Firstly, fuzzy theory can eliminate uncertainty caused by various senors. Secondly, a combination of the K2 algorithm and genetic algorithm is capable of learning the structure and parameter of the Bayesian network via data. Thirdly, bagging learning focuses on exploiting the implicit dependency of features to increase the accuracy and improve the ability of generalization. Experiments show that the proposed RBNF algorithm has a better performance than others.

References

1. Ma, J., Liu, L.: Sea-battlefield situation assessment based on a new method combining dynamic Bayesian network with pattern matching. In: Proceedings of 2014 IEEE Chinese Guidance, Navigation and Control Conference, pp. 1764–1769 (2014)
2. Changliang, X., Yuhui, W., Wenlei, A.: Situation assessment in the warships-airplanes joint operation based on parameter learning in bayesian network. In: Proceedings of 2014 IEEE Chinese Guidance, Navigation and Control Conference (2014)
3. Zihao, L., Bin, Z., Ning, Z., et al.: Hierarchical network threat situation assessment method for DDoS based on D-S evidence theory. In: IEEE International Conference on Intelligence & Security Informatics. IEEE (2017)
4. Li, X., Li, X., Zhao, Z.: Combining deep learning with rough set analysis: a model of cyberspace situational awareness. In: International Conference on Electronics Information & Emergency Communication. IEEE (2016)
5. Park, S., Sohn, M., Jin, H., et al.: Situation reasoning framework for the Internet of Things environments using deep learning results. In: IEEE International Conference on Knowledge Engineering & Applications. IEEE (2017)
6. Li, C., Cao, M., Tian, L.: Situation assessment approach based on a hierarchic multi-timescale Bayesian network. In: International Conference on Information Science & Control Engineering. IEEE (2015)
7. Golestan, K., Khaleghi, B., Karray, F., et al.: Attention assist: a high-level information fusion framework for situation and threat assessment in vehicular ad hoc networks. J. IEEE Trans. Intell. Transp. Syst. **17**(5), 1271–1285 (2015)
8. Cooper, G.F., Herskovits, E.: A Bayesian method for the induction of probabilistic networks from data. J. Mach. Learn. **9**(4), 309–347 (1992)
9. Laskey, K.B.: MEBN: A Language for First-Order Bayesian Knowledge Bases. Elsevier Science Publishers Ltd. (2008)
10. Wu, Y.F., Ng, S.C.: Combining neural learners with the Naive Bayes fusion rule for breast tissue classification. In: IEEE Conference on Industrial Electronics & Applications. IEEE (2007)
11. Feng, L.: Classification problem of ANN in assessment of environmental quality. J. Sci. Total Environ. **441**, 111 (2012)
12. Zhang, J., Huang, S., Yu, J., et al.: Study on the voltage qualification rate prediction from the voltage contour situation picture based on support vector machine. In: Chinese Control Conference. IEEE (2017)

Target Tracking with the Prediction of Trajectory and Kalman Filter in Wireless Sensor Networks

Xiushe Zhang[1], Qiqi Ding[1], Xiaoquan Hu[1], Qin Liu[1(✉)], and Jianshe Wu[2]

[1] The 20th Research Institute of China Electronics Technology Group Corporation, Xi'an 710068, China
liuqin20s@163.com
[2] The Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, Xidian University, Xi'an 710071, China

Abstract. Wireless sensor networks (WSNs) have become a leading solution in many important applications such as monitoring, target tracking and so on. Traditional target tracking algorithms rarely take the obtained optimal state value at the previous moment into account. In this paper, we present a new target tracking method named the prediction of trajectory (PT) which is based on the trajectory of the target. The formulas are deduced at first. Then two improved Kalman Filter (KF) based on PT are highlighted. In addition, the new methods have been independently used to introduce the PT to the extend Kalman Filter (EKF) and the Maximum Likelihood and Kalman Filter (MLKF), respectively. Simulation results show that the new improved algorithms based on PT perform better than other simple methods.

Keywords: Wireless sensor networks · Target tracking · Trajectory · EKF · Prediction of trajectory

1 Introduction

Today, localization based applications for wireless systems have attracted more and more attentions [1]. This interest is expected to grow further with the development of WSNs applications [2]. A WSN consists of some small sensors which contain limited resources [3]. Target tracking is an important application in which target should be localized and tracked as it moves in the surveillance field. Target tracking is widely used in military, industry, agriculture, forestry, animal husbandry and other fields [4].

Most algorithms to track and localize target are based on different types of measurements, e.g. angle of arrival (AOA), time of arrival (TOA), received signal strength (RSS) [4]. The distance measurement is among the classical types in the target tracking, and some distance-based algorithms have been proposed recently [5], in which the Kalman filter (KF) is the most important algorithm. However, a significant drawback of the KF is that the result of positioning may seriously depends on the actual state of target and ignores the trajectory information.

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 403–409, 2020.
https://doi.org/10.1007/978-981-32-9244-4_58

This paper presents a target tracking algorithm named the prediction of trajectory (PT) which brings the trajectory of the target for tracking. It assumes that the trajectory of the target in a short time is stable, and the state of the current time is predicted by the obtained optimal state value at the previous moment. The formula of PT is deduced firstly. Secondly, by introducing the PT to the different two phases of KF, we propose two new improved KF methods. Finally, we experiment with the proposed methods by extend Kalman filter (EKF) and Maximum Likelihood and Kalman Filter (MLKF) [3] respectively.

2 Kalman Filtering

The KF is a classic method to solve the target tracking in WSN. The filtering process is divided into two phases: the prediction phase and the estimation phase.

At the first stage (the prediction phase, also called the time update phase), the target motion model is used to update time and predict the target station at time t_{k+1} (the current time at which the target station is unknown), together with the corresponding covariance.

$$\widehat{x}_{k+1|k} = F_k \widehat{x}_{k|k}, \tag{1}$$

$$P_{k+1|k} = F_k P_{k|k} F_k^T + G_k Q_\omega G_k^T, \tag{2}$$

where F_k is state transition matrix. Q_w is the covariance matrix of acceleration. $\widehat{x}_{k|k}$ is the known target state estimate through the last iteration at time t_k, $\widehat{x}_{k+1|k}$ is the predicted target station value at time t_{k+1}, $P_{k|k}$ and $P_{k+1|k+1}$ is the covariance matrix value corresponding to the $\widehat{x}_{k|k}$ and $\widehat{x}_{k+1|k}$, respectively.

Before the second stage, the converted measurement $\overline{z}_{k+1} = [\overline{x}(k+1) \quad \overline{y}(k+1)]^T$ is needed to be considered, which is useful when filtering. We use the following linear representation in the target state:

$$\overline{z}_{k+1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} x_{k+1} + v_{k+1} = C x_{k+1} + v_{k+1}, \tag{3}$$

where v_{k+1} is the useful converted measurement noise, which is a Gaussian noise with mean zero and covariance R_{k+1}. This is also an important step in extend Kalman filter (EKF). At the second stage (the estimation phase), take the Kalman gain value at time t_{k+1} to be

$$S_{k+1} = C P_{k+1|k} C^T + R_{k+1}, \tag{4}$$

$$K_{k+1} = P_{k+1|k} C^T S_{k+1}^{-1}, \tag{5}$$

where S_{k+1} is a process parameter with no actual meaning, and K_{k+1} is the Kalman gain value at time t_{k+1}. Using the known Kalman gain and measurement at time t_{k+1}, we have the target station and the corresponding covariance matrix value at time t_{k+1}

$$\widehat{x}_{k+1|k+1} = \widehat{x}_{k+1|k} + K_{k+1}\left(\overline{z}_{k+1} - C\widehat{x}_{k+1|k}\right), \tag{6}$$

$$P_{k+1|k+1} = P_{k+1|k} - K_{k+1}S_{k+1}K_{k+1}^{T}, \tag{7}$$

where $\widehat{x}_{k+1|k+1}$ is the estimated state value of target at time t_{k+1}, and $P_{k+1|k+1}$ is the covariance matrix value corresponding to $\widehat{x}_{k+1|k+1}$.

3 Target Tracking Algorithm

In this section, we discuss how to using trajectory for target tracking in detail.

3.1 Prediction of Trajectory

We now discuss what the PT is. For example, let us consider the target motion trajectory in Fig. 1(a), and the trajectory between the two dashed lines is enlarged to the trajectory shown in Fig. 1(b). We assume that node A is the position of target at time t_{k-1}, node B at time t_k and node C at time t_{k+1}. There are some known values shown as position $A(x(k-1), y(k-1))$ and $B(x(k), y(k))$, and velocity $\left(v_x(k), v_y(k)\right)$, respectively. Assuming that the unknown position of node C is $C(x(k+1), y(k+1))$, we have $O(x(k-1)+x(k+1)/2, y(k-1)+y(k+1)/2)$. The slope of the line OB is $t = -v_x(k)/v_y(k)$, and it is defined in the following equation:

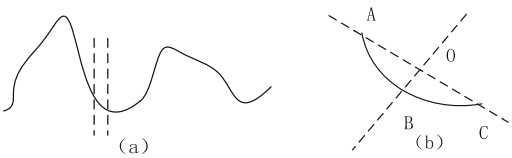


Fig. 1. (**a**) A target motion trajectory, (**b**) A small segment of the trajectory between the two dashed lines from A to B, then to C.

$$y(k) - (y(k+1) + y(k-1))/2 = t(x(k) - 0.5(x(k+1) + x(k-1))), \tag{8}$$

then we have the equation of line AC:

$$y(k+1) - y(k-1) = -\frac{1}{t}(x(k+1) - x(k-1)). \tag{9}$$

Solve the above two equations to get the position of node C as

$$\begin{cases} x(k+1) = x(k-1) - \frac{2t^2(x(k-1)-x(k)) + 2t(y(k)-y(k-1))}{1+t^2} \\ y(k+1) = \frac{(t^2-1)y(k-1) + 2t(x(k-1)-x(k)) + 2y(k)}{1+t^2} \end{cases}, \tag{10}$$

Considering the speed value of node B as the speed of node C, we have the target station value of node C at time t_{k+1}: $a_{k+1} = \left(x_{k+1}, v_x(k), y_{k+1}, v_y(k)\right)$.

3.2 Improved KF Based on PT

In order to take full use of the advantages of PT and KF, we propose two new improved KF methods based on PT.

On the one hand, we introduce the PT to the estimation phase of KF to get the first improved KF method. The specific introduction method is given by (11).

$$\hat{\hat{x}}_{k+1|k+1} = \alpha \hat{x}_{k+1|k+1} + (1\text{-}\alpha)\, a_{k+1}, \tag{11}$$

where $\alpha \in [0,\ 1]$ is a new parameter. The final positioning result at time t_{k+1} is $\hat{\hat{x}}_{k+1|k+1}$ which can be changed by changing the size of α.

The MLKF [3] is an effective algorithm for target tracking in WSN based on the combination of KF and maximum likelihood estimation using distance measurements. It's a typical KF-based target tracking method that seriously depends on the actual state of target and measurements of sensors and ignores the trajectory information. Therefore, the first improved KF method can be introduced to MLKF by adding (11) after the (18) in [3], and we renamed the new algorithm as IMLKF (Improved MLKF).

On the other hand, the KF is divided into two phases, and we can apply the PT to the prediction phase of KF to get the second improved KF method. The specific improvement is given by the following equation:

$$\hat{\hat{x}}_{k+1|k} = \beta \hat{x}_{k+1|k} + (1-\beta)\, a_{k+1} \tag{12}$$

where $\beta \in [0,\ 1]$ is a new parameter. $\hat{\hat{x}}_{k+1|k}$ is the predicted target station value at time t_{k+1} and can be changed by changing the size of β.

The second improved KF method can be applied to EKF by adding (12) after (1) and replacing the $\hat{x}_{k+1|k}$ in (6) with $\hat{\hat{x}}_{k+1|k}$. It can be introduced to MLKF by adding (12) after (16) in [3] and replacing the $\hat{x}_{k+1|k}$ in formula (18) in [3] with $\hat{\hat{x}}_{k+1|k}$. We renamed the two new method GEKF (Generalized EKF) and GMLKF (Generalized MLKF), respectively.

4 Experimental Results

In experiments, Python 3.6 is used to simulate and compare the proposed approaches with the original EKF and MLKF. Parameters' settings are shown in Table 1.

Table 1. General parameter list.

Parameter	Δt_k	μ_n	δ_n^2	μ_γ	δ_γ^2	$\widehat{x}_{0\|0}$	$P_{0\|0}$
Value	0.1 s	0	0.01	0	0.01	$x_0 + 0.1 * random(4,1)$	$0.01I_4$

Assuming n_i and γ_i are the additive and multiplicative Gaussian noises of sensor i with means μ_n and μ_γ and covariances δ_n^2 and δ_γ^2, respectively, Δt_k is the sampling time interval. We use the two-dimensional target motion model, e.g. nearly-constant-velocity linear model (LM) and nearly-constant-speed circular model (CM) [6]. The rectangular area of the sensing area (0 m, 0 m) to (10 m, 10 m) in LM and (0 m, 0 m) to (2 m, 2 m) in CM, respectively. The WSN consists of 100 sensor nodes in LM, which are randomly placed in the radius of both sides of the motion trajectory, and 4 nodes in CM, which are fixed at the four corners.

The detecting radiuses of the sensors are 2 m and 3 m in LM and CM, respectively. The angular velocity is 0.122 rad/s. The initial state is given as $x_0 = [0.0 \quad 1.0 \quad 2.0 \quad 0.3]^T$ and $x_0 = [1.0 \quad 0.05 \quad 0.5 \quad 0.0]^T$ in LM and CM respectively. Each model is simulated 100 times by each method and average results are considered for comparison.

The RMSEs of different tracking approaches are compared in Fig. 2. These experimental results show that the new methods give a better estimate than the two KF-based method, respectively, both in LM and CM. This is because the proposed methods utilize not only the sensing information but also the trajectory information of the moving target. Apparently, the smaller the noise variance is, the more the performance improvement of the two new algorithms. This is because smaller noise variance gives the PT an accurate result, and the fusion results with a smaller α will be more accurate too. When the noise variance is bigger, the original motion model is not accurate. In this case, the EKF method is getting less accurate than method combined PT and EKF directly. As is shown in Table 2, the proposed IMLKF and GMLKF methods perform better than MLKF and runtime reduce by 0.02 s approximately. The RMSE can increase up to 57.21% by adjusting parameters.

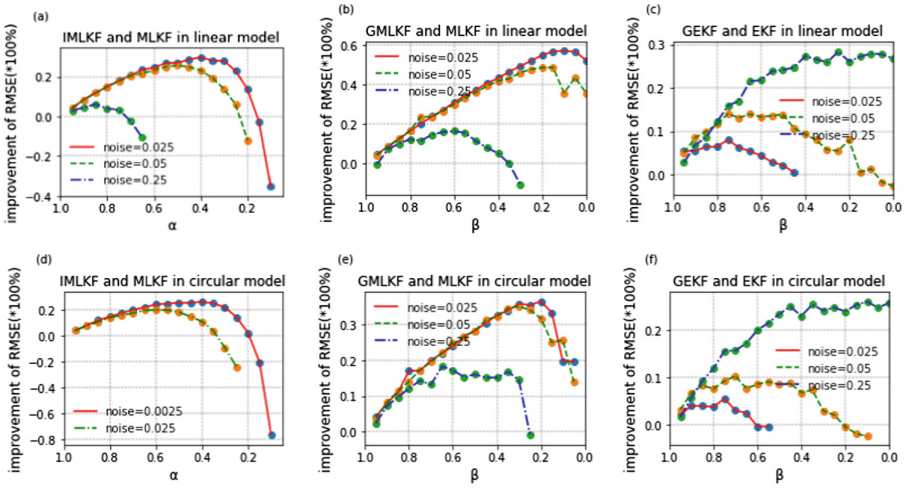


Fig. 2. Comparison between different methods with different noise variance in LM and CM

Table 2. Parameters and running time when the improvement of RMSE is maximized.

Values	LM			CM		
	IMLKF	GMLKF	GEKF	IMLKF	GMLKF	GEKF
Noise variance	0.025	0.025	0.25	0.0025	0.025	0.25
α/β	0.4	0.1	0.1	0.4	0.2	0.1
Improvement %	29.71	57.21	27.93	26.21	36.48	26.05
Time difference	−0.022 s	−0.026 s	0.004 s	0.008 s	−0.238 s	0.008 s

5 Conclusion

In this paper, with the depth analysis of the trajectory and KF-based tracking methods, we propose a new target tracking algorithm named PT, which is based on the trajectory of the target. The PT can be used for both final localization and pre-localization of the target. Therefore, it has been successfully introduced to EKF and MLKF by two methods, respectively, and some simulation and experimental results have shown that the new improved KF based on PT can improve the tracking accuracy compared to the EKF and MLKF, respectively.

References

1. Mahfouz, S., Mourad, C.F., Honeine, P., et al.: Target tracking using machine learning and Kalman filter in wireless sensor networks. IEEE Sens. J. **14**, 3715–3725 (2014)
2. Zhu, G., Zhou, F., Xie, L., et al.: Sequential asynchronous filters for target tracking in wireless sensor networks. IEEE Sens. J. **14**, 3174–3182 (2014)

3. Wang, X., Fu, M., Zhang, H.: Target tracking in wireless sensor networks based on the combination of KF and MLE using distance measurements. IEEE Trans. Mobile Comput. **11**, 567–576 (2012)
4. Cho, H., Kwon, Y.: RSS-based indoor localization with PDR location tracking for wireless sensor networks. AEUE Int. J. Electron. Commun. **70**, 250–256 (2016)
5. Vasuhi, S., Vaidehi, V.: Target tracking using interactive multiple model for wireless sensor network. Inf. Fus. **27**, 41–53 (2016)
6. Misra, S., Singh, S., Khatua, M., Obaidat, M.S.: Extracting mobility pattern from target trajectory in wireless sensor networks. Int. J. Commun. Syst. **28**, 213–230 (2015)

A Personalized Recommendation Algorithm Based on Time Factor and Reading Factor

Xiaoying Zhu, Keda Lu(✉), and Zhenwei Di

School of Data Science and Software Engineering,
Wuzhou University, Wuzhou 543002, China
zhxyl023@qq.com, lukeda@sina.com

Abstract. With the rapid development of the Internet, the data in news, information, education and other application platforms are exploding, which has brought serious information overloading problems to Internet users. The recommendation algorithm is an effective solution to help Internet users select the valuable information from high volume data. Traditional recommendation algorithms ignore the change of users' interest over time and cannot provide users with effective and reasonable recommendation lists. To solve the above problems, this paper proposes a personalized recommendation method that time factor and reading factor into consideration. This method retrieves and compute time factor and reading factor in the process of user tag selection and then produce customized information recommendation list tailored for users. The experiment shows that the accuracy of our algorithm is higher than the traditional recommendation algorithm by considering time factor and reading factor.

Keywords: Personalized recommendation · TF-IDF · Convolutional neural network · Time factor · Reading factor

1 Introduction

With the rapid development of the Internet, the data of news, information, education and other application platforms are exploding, bringing serious information overloading problems to Internet users [1]. Users have higher and higher demands on quick and accurate retrieval of personalized information and resources, which is also a hot research topic that many researchers are interested in.

User's browsing preference is influenced by many factors, and one of the most influential factors is the time factor. Specifically, the hidden features of the expressed content are different for each user at different time. For example, for content that a user is not interested it, the user spends a short time on it, but for content that he is interested in, he spends longer time on it. Therefore, how to mine the potential behavior habits of users to improve the recommendation algorithm is the focus of this paper.

This work was supported by the Guangxi Science and Technology Project (Guike AB16380273) and Guangxi higher education undergraduate course teaching reform project (Guike 2017JGA303) (Guike 2017JGB390).

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 410–417, 2020.
https://doi.org/10.1007/978-981-32-9244-4_59

At present, personalized recommendation algorithms can be classified into three categories content-based collaborative filtering algorithm, user-based collaborative filtering algorithm and hybrid recommendation algorithm. Content-based collaborative filtering algorithm measures the similarity between content by users' comments, and "beer and diapers" is the most typical example. However, at present, most of this algorithm adopts the algorithm of manual extraction, which cannot be extracted normally in large quantities, thus seriously restricting the development of this algorithm [2]. The user-based collaborative filtering algorithm evaluates the similarity between users through their behaviors towards different contents. However, this algorithm only focuses on the degree of users' preference for contents, and the problem of data sparsity occurs in the case of large data volume, thus resulting in the decline of recommendation performance [3]. The traditional hybrid recommendation algorithm recombines the first two recommendation algorithms, but there are still some problems in the hybrid recommendation algorithm, such as data sparsity, cold start, excessive dependence on manual work and so on [4].

In view of the above problems, we proposed a personalized recommendation algorithm that integrates time factor and reading factor. This algorithm firstly considers that the time of browsing information and reading information is related to the selection of user tags. The earlier you browse, the less important it is. The longer you read, the more important it is. Secondly, we construct a model for selecting user tags, which introduces hidden features of users by introducing time factors and reading factors. Finally, we introduce a tag matching algorithm, which considers the dependence of hidden features between user tags and news to construct the relationship between user tags and information, and to recommend information of interest to users. This algorithm considers the influence of time factor and reading factor on user interest distribution, which substantially improves the efficiency and accuracy of recommendation.

2 Problem Statement

The personalized recommendation algorithm uses user's browse records as training data in the process of solving the information overloading problem of news, information and shopping website platform. Traditional recommendation algorithm ignores the characteristics that users' interest may change over time, so it is difficult to provide accurate and comprehensive recommendation lists for users. Therefore, we are thinking about how to effectively use the user's browsing history to improve the accuracy of recommendation. For the information platform, there are several core problems that need to be solved. First, how to obtain users' hidden features from browsing history records; Second, how to build relationships and complete recommendations with the user's hidden features and information.

In response to the specific problem to be solved above, we define the problem as: How do we predict the user's preference for information based on the association between the hidden features and information in the user's browsing history? In order to solve this problem, we need to extract the hidden features of the user, construct the connection between the user and the information through the hidden features of the user, and use this connection to complete the recommendation. Therefore, this paper considers the fusion time factor and reading factor to complete the recommendation.

3 The Personalized Recommendation Algorithm

We propose a recommendation method to mine user's preference for news from different perspectives by acquiring users' historical tags. To better represent the relationship between the various data, we define the following data structure:

Definition 1: The user data set is $User = \{u_1, u_2, \ldots, u_N\}$, where N is the total number of users; The information data set is $Info = \{i_1, i_2, \ldots, i_P\}$, where P is the total number of information; The tag data set is $Tag = \{t_1, t_2, \ldots, t_Q\}$, where Q is the total number of information tags; The browsing time of information is $Time = \{tm_1, tm_2, \ldots, tm_p\}$, the user's reading time of information is expressed by $Read = \{r_1, r_2, \ldots, r_p\}$.

Definition 2: An user hidden interest for different information is denoted with $Tag_{user\_tm} = \{t_1tm_1, t_2tm_2, \ldots, t_Qtm_p\}$, and the hidden interest in reading different information is represented by the set $Tag_{user\_r} = \{t_1r_1, t_2r_2, \ldots, t_Qr_p\}$.

In summary, the issues raised in this paper can be formalized as: $R \leftarrow (Tag_{user}, Tag)$, where R is the similarity calculation result returned, Tag is the Tag set of information resource, and Tag_{user} is the user-specific Tag set.

3.1 Tag Extraction

We apply the TF-IDF algorithm to extract information tags. TF-IDF is a common weighted technique for information retrieval and data mining. TF is the Term Frequency, IDF is the Inverse Document Frequency. The tag of information is extracted from the keywords of the text, and the term frequency of candidate keywords is expressed by the following formula:

$$W_i^{tf} = \frac{n_i}{\sum_{j=1}^{k} n_j} \text{ [5]} \tag{1}$$

where, the W_i is the candidate keywords, and n_i represents the number of occurrences of candidate keywords. The denominator represents the sum of the occurrence times of all k candidate keywords.

IDF inverse text frequency is usually obtained by dividing the number of documents with candidate keywords by the total number of documents and then taking the natural logarithm of the quotient. However, since there is no standard information corpus at present, this paper considers from another perspective that the term frequency of candidate keywords in the corpus is expressed by the following formula:

$$W_i^{idf} = \ln\left(\frac{W_i^s}{W_i^I DF}\right) \text{ [5]} \tag{2}$$

where, the W_i^s is the total term frequency of words with the same length of candidate keyword W_i in the corpus.

The weight of words in the information abstract is higher than the term frequency in the information body, so the position weighting is introduced in the process of weighting. When keyword W_i appears in the information title, it should be denoted with $W_i^{loc} = t_{loc}$, on the contrary, W_i^{loc} is 0.

The weight assignment is expressed as follows according to the length of keywords:

$$W_i^l = \frac{len(W_i)}{\max(len(W_1) \cdots len(W_k))} \text{ [5]} \tag{3}$$

where, the $len(W_i)$ represents the actual length of the candidate keywords, and the denominator represents the length of the longest word in the candidate keywords.

Nouns and phrases with nominal components are given greater weight coefficient, when the words i is a noun word, it is expressed as $W_i^P = t_{pos}$, on the contrary, W_i^P is 0.

The final weighting formula is expressed as follows:

$$\mathrm{W} = W_i^{tf} * W_i^{idf} * \left(1 + W_i^{loc} + W_i^l + W_i^p\right) \text{ [5]} \tag{4}$$

The text keyword set is sorted from high to low according to the word weight, and the first 20 keywords are extracted as the information tag. The tag set is expressed as follows:

$$Tag = \{t1, t2, \ldots, t20\}$$

3.2 Tag Selection

The tag selection is based on the user's historical reading of the information, and the tag of the information is associated with the user. Firstly, the tag is de-weighed, and then the tag of the information is sorted according to the weight, so as to obtain the exclusive tag of the user.

The weight of user tags decreases with time. The time spent browsing information and the time spent reading information are related to user tags, and the importance of browsing decreases with time. Reading increases in importance over time. Therefore, we define the time factor t and the reading factor r, among then time factor t is the attenuation factor and reading factor r is the enhancement factor. The position, length and part of information content keywords determine the weight and content determines the label. After further conversion, the formula is expressed as follows:

$$\mathrm{W} = W_i^{tf} * W_i^{idf} * \left(1 + W_i^{loc} + W_i^l + W_i^p\right) * Tm_i * R_i \tag{5}$$

In the process of user tag selection, we fuse time factor and reading factor, and select 20 user-specific tags according to the weight order:

$$Tag_{user} = \{t1, t2, \ldots, t20\}$$

3.3 Tag Matching

Traditional recommendation methods match tags according to similarity, but they depend heavily on the correctness of tag selection. In this paper, similarity calculation is carried out by using tags of resources [6], and cosine similarity is adopted as the measurement standard.

Resource features are represented by tag-based resource feature vector $\overrightarrow{I_k} = (n_{k1}, n_{k2}, \cdots, n_{ki}, \cdots, n_{kL})$, among which $k = 1, 2, 3, \ldots, N$; n_{ki} represents that the normalized value of t_i is used to mark information i_k.

All information features can be represented by the following information feature vectors:

$$I_{k\times k} = \begin{bmatrix} n_{11} & \cdots & n_{1k} \\ \vdots & \ddots & \vdots \\ n_{k1} & \cdots & n_{kk} \end{bmatrix} [7] \tag{6}$$

Our similarity calculation method uses cosine algorithm:

$$\sin(I_j, I_k) = \cos\left(\overrightarrow{I_J}, \overrightarrow{I_K}\right) = \frac{\overrightarrow{I_J} \cdot \overrightarrow{I_K}}{\left|\overrightarrow{I_J}\right| \times \left|\overrightarrow{I_K}\right|} [7] \tag{7}$$

Among them, Eq. 7 can be expressed as a matrix:

$$s_{p\times p} = \begin{bmatrix} 1 & \cdots & S_{1j} & \cdots & S_{1P} \\ \vdots & \ddots & & & \vdots \\ S_{j1} & & 1 & & S_{jP} \\ \vdots & & & \ddots & \vdots \\ S_{P1} & \cdots & S_{Pj} & \cdots & S_{PP} \end{bmatrix} [7] \tag{8}$$

where S_{ij} represents the similarity between resources i_i and i_j.

We rank the similarity calculation results of user tags and information tags in order, and select the first N information as the recommendation results.

4 Experimental

4.1 Datasets

Our data set comes from the "Internet News Platform" (Wechat Public Number: AI_30s) we created. Data items include information title, information content, publication time, label source, user ID user browsing record user browsing time, etc.

4.2 Evaluation Metric

We use the Precision P and Recall R to evaluate the performance of the recommendation system, where, P is the ratio of the amount of data read by users in the recommendation list to the total amount of data in the recommendation list, R is the proportion of recommended data in the total amount of information data read by users, and F value is the composite index, namely the harmonic average of the Precision and Recall rate. The calculation formulas are as follows:

P $= (Ru \cap Tu)/Ru$, where $(Ru \cap Tu)$ indicates the amount of data read by the user in the recommendation list, and Ru indicates the total amount of data in the recommended list.

R $= (Ru \cap Tu)/Tu$, where $(Ru \cap Tu)$ indicates the amount of data read by the user in the recommendation list, and Tu indicates the total amount of information that the user reads.

$$F = P * R * 2/(P + R)$$

Among them, Ru is a list of information that is recommended for user u, and Tu is a reading list for that user. That is, the recommended cases can be divided into 4 cases, as shown in Table 1:

Table 1. Recommended situation table

System user	Recommendation	Not recommendation
Interest	$Ru \cap Tu$	$\neg Ru \cap Tu$
No interest	$Ru \cap \neg Tu$	$\neg Ru \cap \neg Tu$

4.3 Performance Evaluation

The experimental results are divided into two parts, one is the evaluation of the algorithm on the online recommendation algorithm, and the other is the user recall rate of the "Internet News Platform" (WeChat public number: AI_30s) actively recommending popular information. Based on the comparison of the experimental results of the above two parts, we obtain the traditional recommendation algorithm and the P value,

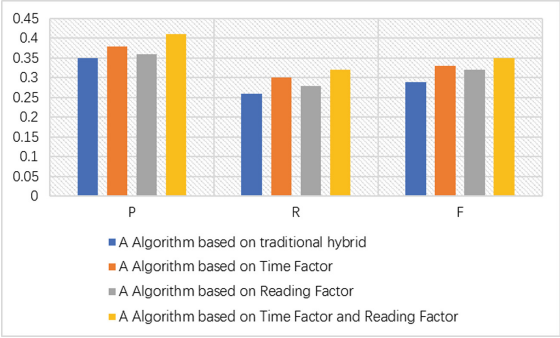


Fig. 1. Comparison of recommendation indexes

R recall rate and F value of the recommendation algorithm that integrates the time factor and the reading factor based on the traditional recommendation algorithm. Figure 1 is a comprehensive experiment result.

From the experimental results, the traditional hybrid recommendation algorithm has the lowest accuracy rate P, recall rate R and F value of the recommendation results recommended by the user, which are 0.35, 0.26, and 0.29, respectively, algorithm that take both time factor and reading factor into consideration. The recommendation algorithm that only considers the time factor or the reading factor is slightly higher than the traditional hybrid recommendation algorithm, and the fusion time factor and the reading factor has the highest accuracy rate P, recall rate R and F. they are 17.1%, 23.0%, and 20.6% higher than the traditional hybrid recommendation algorithm. In summary, it can be concluded that the integration of time factor and reading factor in the process of recommending information to users can greatly improve the accuracy of recommendation results and provide users with more personalized recommendations.

5 Conclusions

Personalized recommendation is widely used in many fields, and it has very important practical significance. The accuracy of a personalized recommendation algorithm is affected by many aspects. The browsing time and reading time of users are potentially related to the user's hobbies. In view of the above problems, this paper proposes a recommendation algorithm that can construct the connection between users and information according to the hidden features in the user's browsing record. The method first constructs a model for selecting user tags, then introduces a tag matching algorithm, and finally constructs the association between user tags and information, and completes recommending information of interest to users. Through a large number of experiments, this paper proves that the recommended method with time factor and reading factor is more aligned with the user's reading interest, and the accuracy of recommendation has been significantly improved.

The recommended method of fusion time factor and reading factor proposed in this paper still has many problems to be solved in the practical application, such as the impact of national statutory holidays on the recommendation results. If the national legal holidays can be integrated into the recommendation system, it will be the recommended accuracy is higher, which is the next step in this paper.

References

1. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Trans. Knowl. Data Eng. **17**(6), 734–749 (2005)
2. Wang, S., Tang, J., Wang, Y., Liu, H.: Exploring hierarchical structures for recommender systems. IEEE Trans. Knowl. Data Eng. **30**(6), 1022–1035 (2018)
3. Shi, Y., Larson, M., Hanjalic, A.: Collaborative filtering beyond the user-item matrix: a survey of the state of the art and future challenges. ACM Comput. Surv. **47**(1), 1–45 (2014)

4. Mooney, R.J., Roy, L.: Content-based book recommending using learning for text categorization. In: Proceedings of the Fifth ACM Conference on Digital Libraries, pp. 195–204. ACM (2000)
5. Tf-idf optimization algorithm for news tag extraction version 1.0 – based on the segmentation of jieba. https://blog.csdn.net/say_c_box/article/details/72765014
6. Li, L., Li, C., Chen, H., et al.: MapReduce-based SimRank computation and its application in social recommender system. In: Proceedings of the International Congress on Big Data, Santa Clara, USA, pp. 133–140 (2013)
7. Zhu, X., Cai, Q., Bai, I., et al.: A parallel recommendation algorithm based on tag and collaborative filtering is proposed. High Technol. Lett. **25**(3), 307–312 (2015)

Reliable Blockchain Based Stochastic Game System

Jeong Hoon Jo and Jong Hyuk Park(✉)

Department of Computer Science and Engineering,
Seoul National University of Science and Technology (SeoulTech),
Seoul 01811, Korea
{jojeong3766, jhparkl}@seoultech.ac.kr

Abstract. Recently, due to the development of network and the development of hardware, users can enjoy games anywhere, and as a result various games have been developed so that items in the game using cash and game cache have skyrocketed. The user (participant) uses his cash and game cache to estimate the probability and make the investment. However, there is a problem that the item does not come out even though the number of times of approaching the probability is reached. The reason is that developers can manipulate the probability information stored in the database, and the data of the probability game can be recorded uncertainly so that users can suffer damage. In this paper, we propose a reliable system based on a blockchain for a random item in a game. The developer marks the probability in the blockchain. The user can verify the information stored in the blockchain. Since the event that the user participates in is also stored in the blockchain, the user verifies/checks the information stored in the blockchain, thereby enhancing the reliability of the corresponding game.

Keywords: Blockchain · Stochastic game

1 Introduction

The recent development of wired and wireless networks and the development of hardware that allows games to be played, users can enjoy games regardless of location and time. Developers develop various games to satisfy users' needs. Developers needed funds to maintain their games, and developers have made many attempts to cover them. To make a purchase of a game, to purchase a piece of the update at the time of an additional update, or to sell an item in a game for a fee. Among them, the items in the game occupied a lot of demand, and the developers applied the probability to the items in the game to induce the user to make a lot of demand. The user (participant) uses his cash and game cache to estimate the probability of the item and make an investment. However, there is a problem that the item does not come out even though the number of times of approaching the probability is reached. The reason is that developers can manipulate the probability information stored in the database, and the data of the probability game can be recorded uncertainly so that users can suffer damage. In this paper, we propose a reliable system based on a blockchain for a random item in a game. The developer marks the probability in the blockchain. The user can (participant)

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 418–421, 2020.
https://doi.org/10.1007/978-981-32-9244-4_60

verifies the information stored in the blockchain. Since the event that the user participates in is also stored in the blockchain, the user (participant) verifies/checks the information stored in the blockchain, thereby enhancing the reliability of the corresponding game.

2 Related Work

2.1 Game Market

The game market is experiencing explosive growth with the addition of traditional PC, video and mobile gaming environments due to advances in network and hardware. Mobile cloud enabled us to implement high-performance mobile games in our mobile environment [1], and PC and video games are now available thanks to the development of GPU. Table 1 lists game revenues worldwide. China spends about 37,945 M dollars on games, followed by the United States and Japan [2]. As such, the game market represents a large industry. Game companies make consumers pay the right price for maintaining the game, and the first is that they pay for the game when they buy the game. Second, every time a game is updated, consumers pay for the update. Third, there is a different way of paying the price depending on the amount of game users use. Finally, there is a way to make profits by selling paid items in games.

Table 1. Game market consumption ranking

Rank	Country	Region	Internet population	Total revenues in US dollars
1	China	Asia	850 M	37,945 M
2	United States	North America	265 M	30,411 M
3	Japan	Asia	121 M	19,231 M
4	Republic of Korea	Asia	48 M	5,647 M
5	Germany	Western Europe	76 M	4,687 M
6	United Kingdom	Western Europe	64 M	4,453 M
7	France	Western Europe	58 M	3,131 M
8	Canada	North America	34 M	2,303 M
9	Spain	Western Europe	39 M	2,032 M
10	Italy	Western Europe	40 M	2,017 M

2.2 Blockchain Technology

A blockchain is linked to a chain based on a technique for creating a block that stores transactions. It is stored in a distributed data store. Any public blockchain can read transaction data and add transactions. The data stored in the blockchain is verified from other nodes after the block is created, and the recorded transaction cannot be deleted. Therefore, it is possible to prevent the forgery and alteration of the data, If the block is not occupied by more than 51%, the block will be reliable [4]. There are public

blockchains, private blockchains, and consortium as types of blockchains. In the public blockchain, the bitcoin and ethereum are typically used, and transaction details use a proof-of-work method to verify the modulation by an attacker and to detect and generate an abnormal block [3]. The private blockchain is available only to a specific user, and if it is not possible to disclose the data, it will use the private bloc chain. The private blockchain may not focus on cryptography, and instead uses a competing algorithm such as the Federated Byzantine agreement (FBA) instead of the PoW to validate in the public blockchain [5]. It also has fast processing speed and high data throughput, so it can have fast response speed [6]. Consortium blockchains are only available to consortium users, and data write and read permissions and block creation are only available to consortium participants. Because it has the advantages of public blockchain and private blockchain, it is mainly used between government and private enterprise, it has some anonymity, and it can have fast processing speed [7] (Table 2).

Table 2. Blockchain type

Type	Public blockchain	Consortium blockchain	Private blockchain
Operator	None	Consortium member	Specific operating entity
Read, Write data permissions	Blockchain participants	Consortium member Blockchain participants	Specific operating entity Blockchain participants
Block creation	Blockchain participants	Consortium member	Operator
Block generation algorithm	Proof-of-Work	Practical BFT	Proof-of-Auth
Throughput	Low	High	High
Latency	High	Low	Low
Examples	Bitcoin, Ethereum	R3 Corda	Self-development

3 Conclusion

Recently, the game market has grown rapidly due to the development of network and hardware, and various game services have been operated in a form in which users pay for the maintenance of their games. Players are paying more money from users by adding the odds on the items in the game and the items in the game. In this paper, we propose blockchain based random item acquisition to provide reliable game service. The proposed blockchain based service stores the intrinsic probability of an item in a blockchain. When a user (participant) participates in acquiring a specific item, the user's information is also stored in the blockchain transaction, And the probability of manipulation of the developer can be prevented. Future research will implement a service framework for preventing probability data in other environments.

Acknowledgement. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2016R1A2B4011069).

References

1. Cardellini, V., et al.: A game-theoretic approach to computation offloading in mobile cloud computing. Math. Program. **157**(2), 421–449 (2016)
2. https://newzoo.com/insights/rankings/top-100-countries-by-game-revenues/. Accessed 29 Oct 2018
3. Nakamoto, S.: Bitcoin: a peer-to-peer electronic cash system (2008)
4. Eyal, I., Sirer, E.G.: Majority is not enough: bitcoin mining is vulnerable. In: International Conference on Financial Cryptography and Data Security, pp. 436–454. Springer (2014)
5. Mazieres, D.: The stellar consensus protocol: a federated model for internet-level consensus. Stellar Development Foundation (2015)
6. https://www.ibm.com/blockchain/hyperledger. Accessed 8 June 2018
7. Zheng, Z., et al.: An overview of blockchain technology: architecture, consensus, and future trends. In: 2017 IEEE International Congress on Big Data (BigData Congress). IEEE (2017)

Author Index

J. J. Park et al. (Eds.): MUE 2019/FutureTech 2019, LNEE 590, pp. 423–425, 2020.
https://doi.org/10.1007/978-981-32-9244-4